VOLUME 103 PART B NUMBER 12 NOVEMBER 1956

The Proceedings

OF

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871: INCORPORATED BY ROYAL CHARTER 1921

PART B

RADIO AND ELECTRONIC ENGINEERING
(INCLUDING COMMUNICATION ENGINEERING)

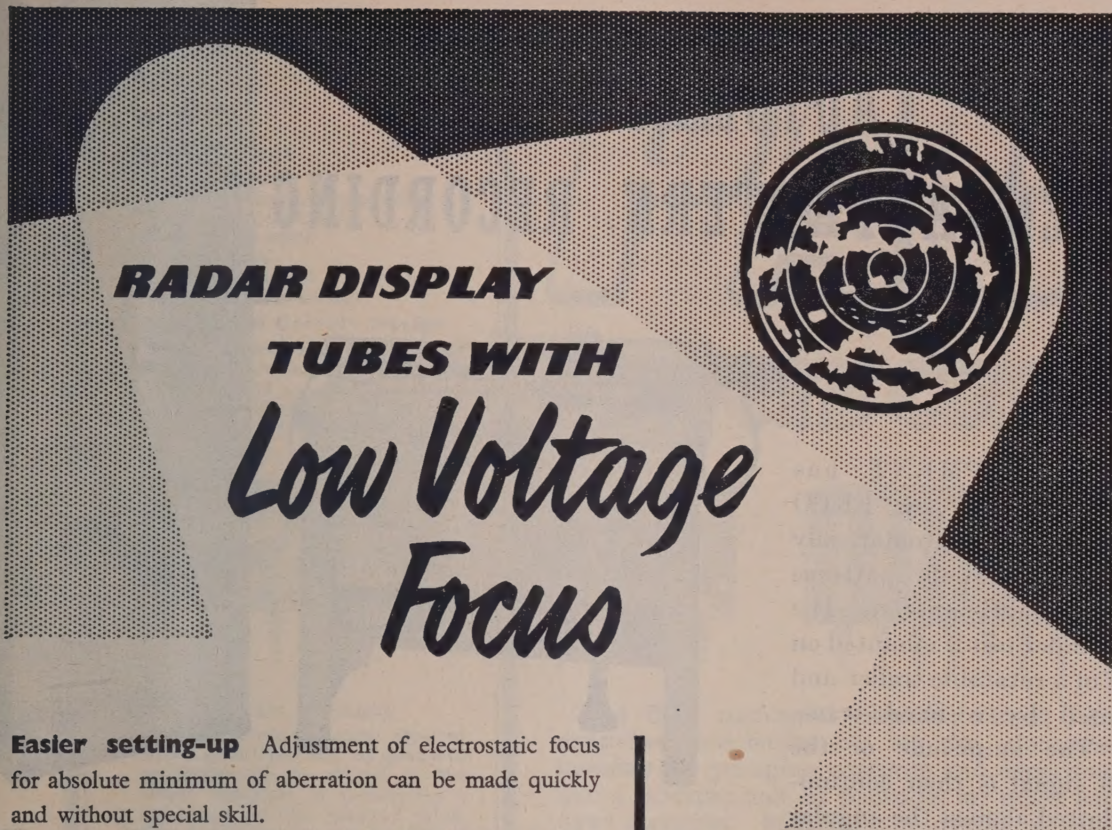
SAVOY PLACE . LONDON W.C. 2

Price Seven Shillings and Sixpence

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871 INCORPORATED BY ROYAL CHARTER 1921

PATRON: HER MAJESTY THE QUEEN



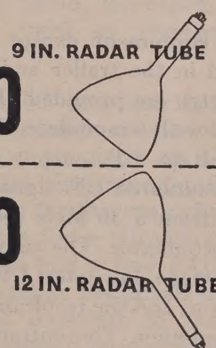
Easier setting-up Adjustment of electrostatic focus for absolute minimum of aberration can be made quickly and without special skill.

Simpler E.H.T. The focus voltage swing required is only ±200V about cathode—e.h.t. units need no longer be loaded with the current wasting potential dividers associated with earlier electrostatic focus tubes.

Lower cost No focus magnet or coil is required. Ordinary carbon track potentiometer across normal low-voltage supply is all that is needed to achieve fine focus.

Space saved Focus at ordinary h.t. potential means that e.h.t. generators can be smaller, and bulky high voltage potential dividers eliminated.

AL22-10



AL 31-10

ABRIDGED DATA FOR BOTH TUBES				
	Typical Operating Conditions			
Focus: Electrostatic Deflection: Magnetic Heater: 6.3V, 300mA Phosphor: long persistence	V_{a2+a4} (kV)	V_{a1} (V)	$-V_g$ for cut-off (V)	V_{a3} focus (V)
	12	300	30 to 70	±200

Write to the address below for full details of these and other cathode ray tubes in the Mullard range.

Mullard

MULLARD LTD., COMMUNICATIONS & INDUSTRIAL VALVE DEPT., CENTURY HOUSE, SHAFTESBURY AVENUE, LONDON, W.C.2

MVT 187

AUTOMATIC ANTENNA PATTERN RECORDING

THIS EQUIPMENT has been developed by EKCO Electronics to automatically record the radiation patterns of any centimetric antenna. The antenna under test is mounted on the roof of a rotatable trailer and illuminated by a fixed transmitter. The amplitude of the received signal is then continuously plotted against the angular traverse of the trailer.

EKCO ANTENNA PATTERN RECORDER type E59

All the equipment except the transmitter unit is mounted in the trailer and remote controls for the transmitter are provided. The received C.W. signal is mixed with a modulated local oscillator signal and the resultant I.F. output combined with an anti-phase modulated I.F. signal. The reference signal is derived from a 30 Mc/s oscillator and servo-driven piston attenuator. The combined signals are fed via a seven-stage, low noise I.F. amplifier to a balanced modulator, and the resultant error signal applied to a servo amplifier. The output of this amplifier drives a servo motor which moves the piston attenuator in such a direction as to reduce the difference between the reference signal and the received signal. A pen attached to the piston drive mechanism records the amplitude of the received signal in terms of the attenuation law of the standard piston which is directly calibrated in dB.

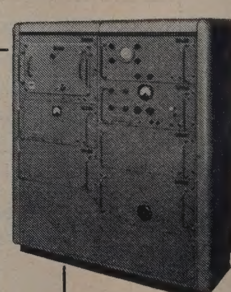
Facilities are available for plotting either on Cartesian or polar co-ordinate graph paper. Piston Attenuators can be supplied to provide amplitude scales of either or both 5 and 10 bB per inch.

The maximum travel is 35 or 65 dB respectively. The Cartesian co-ordinate paper can be run at rates accurately corresponding to two or five degrees per inch.

This equipment is also available to special order for installation in a permanent location

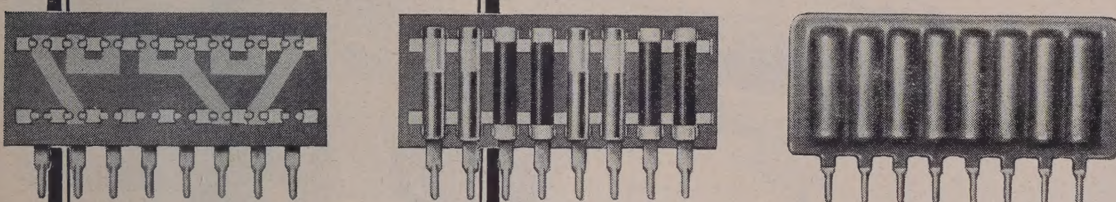
EKCO
electronics

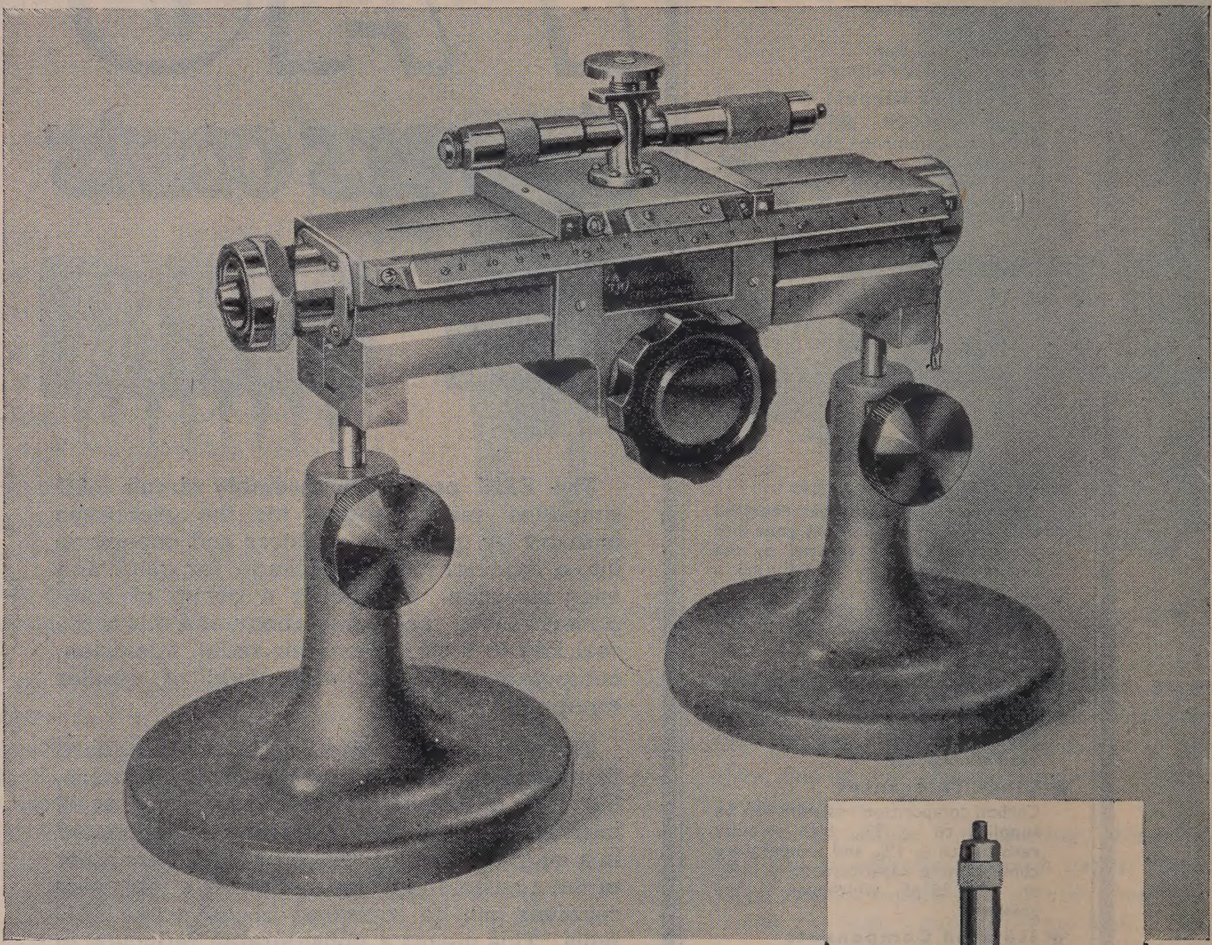
The Mixer section is a plug-in unit and 'X' and 'S' Band versions are available covering the ranges 8250-10,000 Mc/s and 2500-3300 Mc/s respectively. Other frequency ranges can be covered, to special order. The Recorder can be supplied in two forms with either a single 6 ft. (shown above with plotting table) or a twin 4 ft. console rack shown at right.

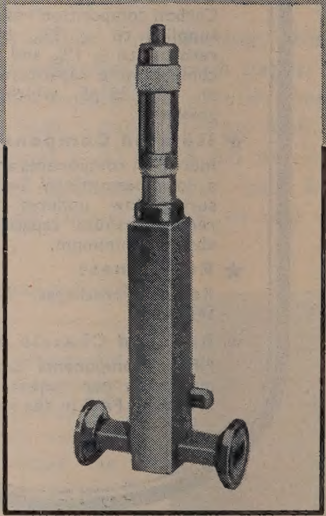


We shall be pleased to discuss this equipment with you

EKCO ELECTRONICS LTD · EKCO WORKS · SOUTHEND-ON-SEA · ESSEX

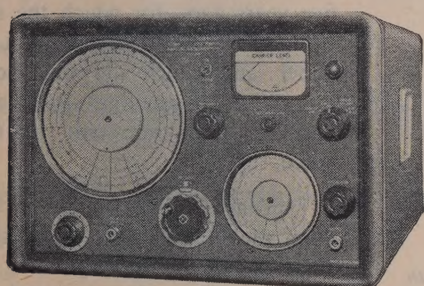






For Precision Measurement up to 4,000 Mc/s . . .

U.H.F. & S.H.F. SIGNAL GENERATOR
TYPE TF 1058



U.H.F. SIGNAL GENERATOR
TYPE TF 1060

Marconi U.H.F. Signal Generator
TYPE TF 1060

Marconi U.H.F. & S.H.F. Signal Generator
TYPE TF 1058

These two Marconi Signal Generators simplify a wide variety of measurements in the ranges 450 to 1250 and 1700 to 4000 Mc/s. The range of each is covered in one band and frequency is directly indicated on a large, clear dial; incremental tuning facilities coupled with inherent high stability provide for the most exacting bandwidth measurements. The wide range of output levels, directly indicated on the dial of a piston attenuator, is suitable for a diversity of applications extending from slotted-line measurements to the testing of high-sensitivity receivers.

ABRIDGED SPECIFICATIONS

TF 1060 *Frequency Range:* 450 to 1250 Mc/s. *Incremental Tuning:* 20 to 200 kc/s per division, with interpolation by fine tuning control. *Frequency Stability:* Better than 0·005% per 10-minute period. *R.F. Output:* 0·15μV to 445mV; also higher outputs up to approximately 3 volts. *Output Impedance:* 50 ohms. *Modulation:* Internal sine and external pulse a.m.

TF 1058 *Frequency Range:* 1700 to 4000 Mc/s. *Incremental Tuning:* 60 to 350 kc/s per division, with interpolation by fine tuning control. *Frequency Stability:* Better than 0·001% per 10-minute period. *R.F. Output:* 0·1μV to 445mV; also higher outputs up to approximately 3 volts. *Output Impedance:* 50 ohms. *Modulation:* Internal square a.m.; external f.m. and pulse a.m.

MARCONI INSTRUMENTS

AM & FM SIGNAL GENERATORS · AUDIO & VIDEO OSCILLATORS · VALVE VOLTMETERS · POWER METERS Q METERS · BRIDGES · WAVE ANALYSERS · FREQUENCY STANDARDS · WAVEMETERS · TELEVISION AND RADAR TEST EQUIPMENT · AND SPECIAL TYPES FOR THE ARMED FORCES

MARCONI INSTRUMENTS LTD · ST. ALBANS · HERTFORDSHIRE · TELEPHONE: ST. ALBANS 56161

London and the South: Marconi House, Strand, London, W.C.2. Tel: COVent Garden 1234

Midlands: Marconi House, 24 The Parade, Leamington Spa. Tel: 1408 *North:* 30 Albion Street, Kingston-upon-Hull Tel: Hull Central 16347

WORLD-WIDE REPRESENTATION

TC 88

latest developments in EQUIPMENT

TRANSISTORISED EQUIPMENT

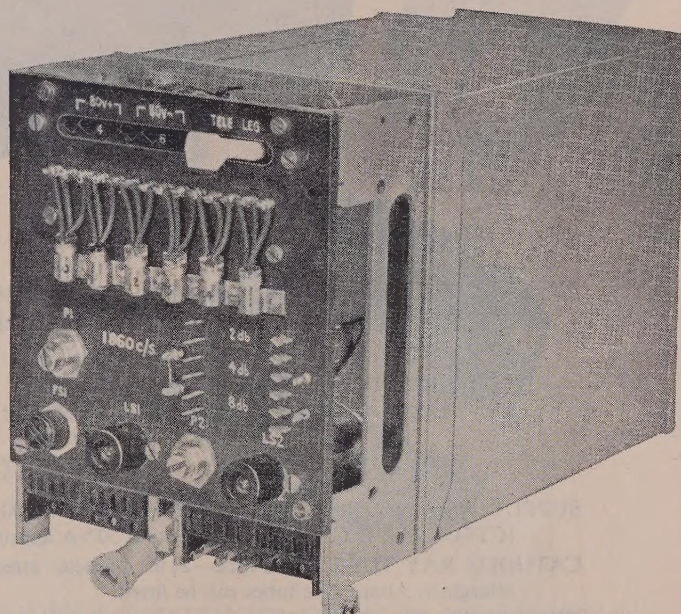
The use of transistors in transmission equipment has many advantages. For example:

The power consumed is extremely low.

The physical size is small.

The heat dissipation is negligible.

The G.E.C. is producing completely transistorised equipment for the following applications:



A complete VF telegraph unit including transmitter and receiver.

RURAL CARRIER SYSTEM

This enables up to ten circuits to be transmitted over one pair of wires, with facilities for terminating one or more circuits at intermediate points. Full information on this equipment is given in standard specification SPO 1030.

VOICE-FREQUENCY TELEGRAPH EQUIPMENT

This extremely compact equipment uses transistors throughout, and operates from a 24-volt DC supply. The system employs frequency shift modulation to provide 24-duplex telegraph channels operating at a modulation rate of 50 bauds over any four-wire speech circuit that effectively transmits frequencies between 300 c/s and 3,400 c/s. A complete terminal is mounted on a single-sided rack 9 ft. high × 1 ft. 8½ in. wide. Full information regarding the equipment is contained in standard specification SPO 1403.

4-WIRE AUDIO AMPLIFIER

This has a maximum gain of 30dB and is capable of delivering a maximum output of +16dBm. The gain frequency distortion over the range 300 c/s to 6 kc/s does not exceed ·5dB relative to 800 c/s.

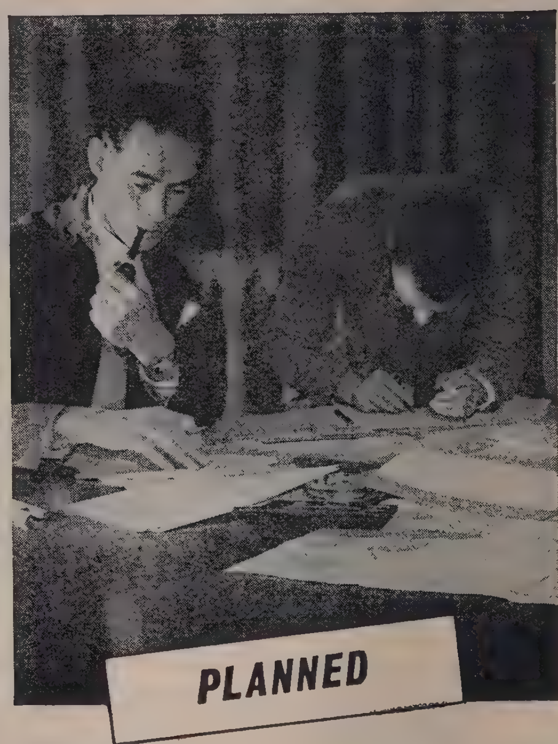
NEGATIVE IMPEDANCE REPEATERS

These are of the shunt and series types for use on loaded 2-wire audio cables.

THE GENERAL ELECTRIC COMPANY LIMITED OF ENGLAND

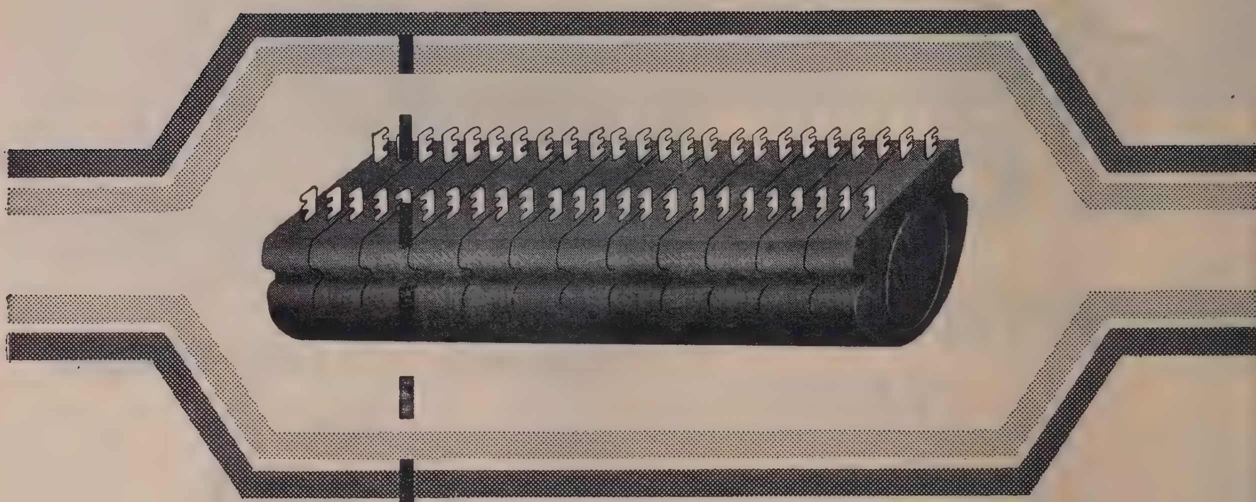
TELEPHONE, RADIO AND TELEVISION WORKS · COVENTRY · ENGLAND

MARCONI



LONG-DISTANCE H.F. TELEGRAPH SYSTEMS High Frequency systems form a major part of world-wide radio telegraph communication services. Marconi's have recently designed new equipment for such systems incorporating the latest electronic developments to save time and labour, reduce operating costs and eliminate faults. The company is unique in the resourcefulness and skill it can bring to the complete engineering of a system from the surveying stage onwards to the maintenance after it has been installed, and the training of the staff to operate it at maximum efficiency.

Loading coils **inside** the cable splice

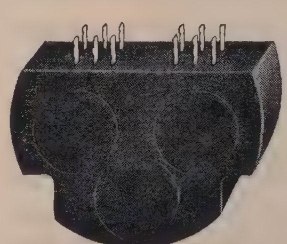


The advantages of the splice loading technique are particularly marked in the loading of small cables of up to 74 pairs. The coils can be included in a jointing sleeve or unit of only slightly larger diameter than would normally be used.

The loading coils in the Mullard L.160 Series are designed specifically for this technique. They are cast in resin, which provides complete protection from climatic conditions and allows a telephone administration to store them ready for building into loading units as and when required. Both single and triple assemblies are available for different sizes of cable.

Ferroxcube pot cores give these coils certain electrical advantages over conventional types, particularly in the loading of higher frequency circuits such as those encountered in programme and carrier applications.

You are invited to write for leaflets describing the Mullard L.160 Series coils and simple units for pole and splice loading.



TRIPLE COIL ASSEMBLY

Mullard Mullard

SPECIALISED ELECTRONIC EQUIPMENT

MULLARD LTD · EQUIPMENT DIVISION
CENTURY HOUSE · SHAFTESBURY AVENUE · W.C.2

(MI453A)

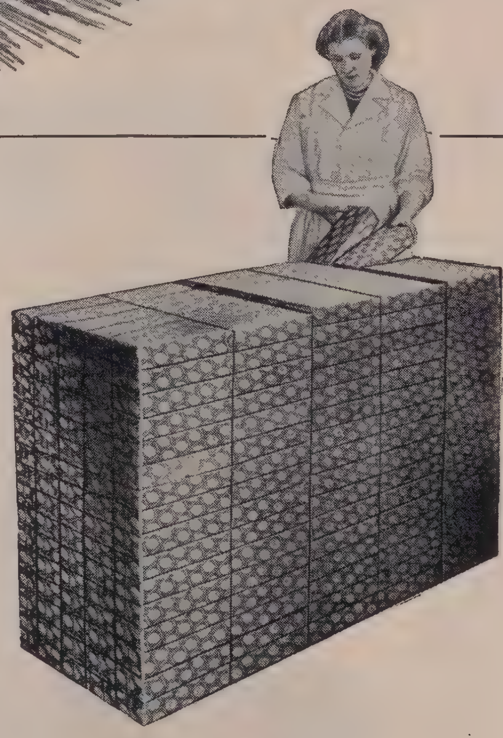
DUBILIER

box clever

with four new autopacks

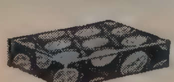
for RESISTORS

DUBILIER AUTOPACKS—designed primarily for loading hoppers for automatic feed systems—also solve the spiky problem of resistor storage. The resistors* are housed in packs of 50, 100, 500 and 2,500 for type BTS, and in packs of 100 and 500 for type BTB—all lined up with connecting wires dead straight, ready for immediate use—and taking up very little space in the process. The large *and* the small user would be well advised to take stock of these handy, space and time-saving packs.



50

100



500

2500

No storage problems here!

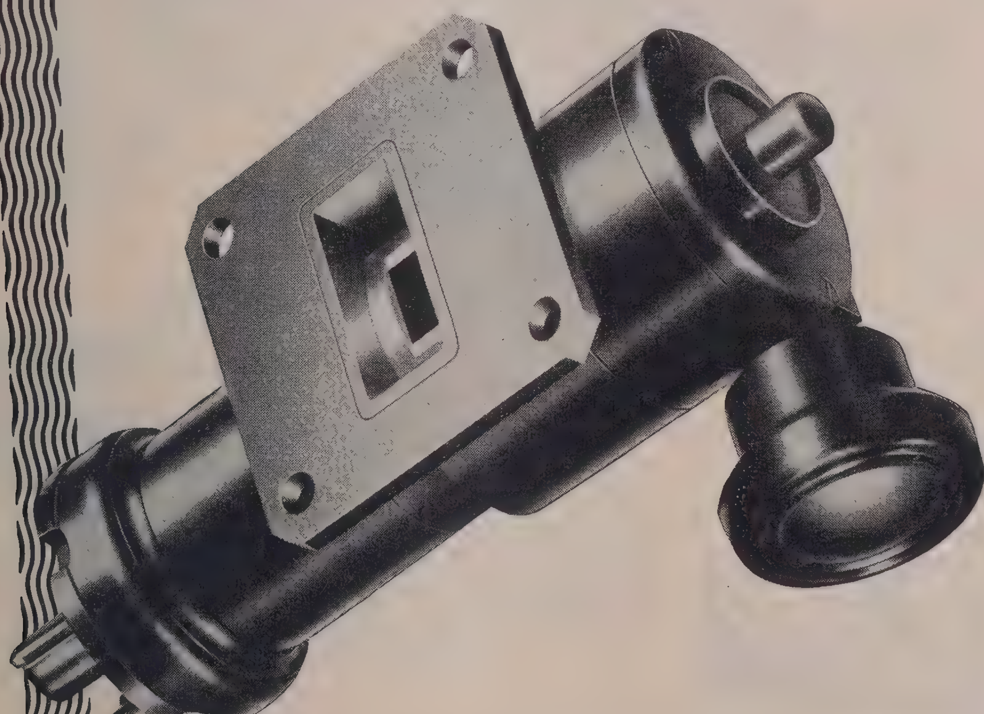
This modest sized stack contains no fewer than one million separate resistors. A pack of 50 would occupy no more room than 10 cigarettes on the workbench. Whether you use resistors in tens, hundreds or thousands—see Dubilier about these new space-saving Autopacks today.

**The resistors are available in two ratings: BTS ½ watt, BTB 1 watt at 70°C. Resistance range is 100 ohms to 10 megohms (BTS) and 390 ohms to 22 megohms (BTB). Each type is completely protected by a phenolic resin housing which is sealed at the ends.*

DUBILIER CONDENSER CO. (1925) Ltd., Ducon Works, Victoria Rd., North Acton, W.3. Phone: ACOrn 2241. Grams: Hivoltcon, Wesphone, London.

DN 173

b



Available for the first time

a complete range of

transistors

—for the radio

and electronics

industry

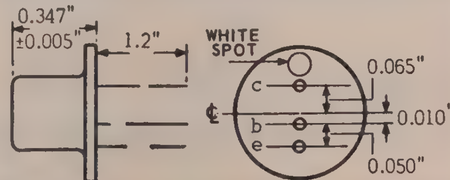
INTRODUCING
the New Metal Clad "Top Hat" TRANSISTORS

★ Hermetically sealed, welded metal-to-metal container insures against moisture penetration even in high humidity at high ambient temperatures.

★ Revolutionary new all-metal construction gives far greater heat dissipation.

★ A complete range of basic samples with mean average characteristics are available, including matched pairs for push-pull output, for Design Engineers in the radio and electronics industry.

TYPE	APPLICATION
XA101	I.F. Amplifier
XA102	Frequency Changer/Oscillator (cut-off frequency 7 Mc/s)
XB102	Intermediate L.F. Stage
XB103	Intermediate L.F. Stage
XC100 Series	Output Stage, including matched pairs for push-pull



Set Manufacturers and other Trade users are invited to send for range of basic samples and technical data.

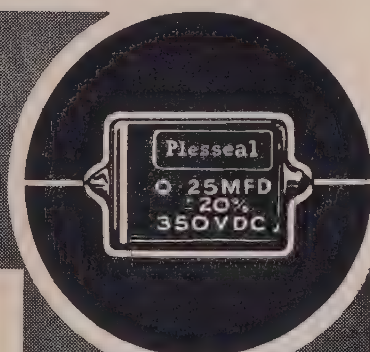
EDISWAN
MAZDA

RV 29

RADIO DIVISION
THE EDISON SWAN ELECTRIC CO. LTD., 155 CHARING CROSS ROAD, LONDON, W.C.2
Telephone: Gerrard 8660 *Member of the A.E.I. Group of Companies* Telegrams: Ediswan, Westcent, London

Plesseal

Climate-proof paper capacitors



There has long been a need for a superior type of paper capacitor—one which is completely reliable and able to withstand every extreme of climatic variation. Plesseal meets this need. Its flattened shape enables greater capacitance with higher working voltage to be contained in a given volume, while saving space on group and tag-panels.

A hard lacquer final coating completely shrouds the capacitor and gives physical protection to the special hard wax housing, thus completely excluding moisture.

Plesseal exceeds the requirement of specification RCS11 Grade H2 for tropical storage and is therefore comparable in this respect to many metal cased paper capacitors, with the additional advantage of costing much less.

-shape new standards in capacitor performance

note these Plesseal features

* **Improved shape saves space**
* **Capacitances from .0005μF to 1μF, in popular and logarithmic values.**
* **Voltage ranges, 150, 250, 350, 500, 750, 100 volts D.C.**
* **Insulation resistance greater than 4000 M Ω.μF at 20°C.**
* **Non-inductive.**
* **Breakdown voltage in excess of 6 times working voltage.**
* **Complete protection against all conditions of humidity.**
* **Working temperature up to 85°C.**
* **Electrostatic screening.**

Plessey

Design Engineers are invited to ask for further details

THE PLESSEY COMPANY LIMITED · KEMBREY STREET · SWINDON · WILTS · Telephone: SWINDON 5461

SC1a

THE PROCEEDINGS OF
THE INSTITUTION OF ELECTRICAL ENGINEERS

EDITED UNDER THE SUPERINTENDENCE OF W. K. BRASHER, C.B.E., M.A., M.I.E.E., SECRETARY

VOL. 103. PART B. No. 12 NOVEMBER 1956

Paper No. 1989 M
Feb. 1956

521.317.7.082.64 : 621.316.89

THE INDIRECTLY HEATED THERMISTOR AS A PRECISE A.C.-D.C. TRANSFER DEVICE

By F. C. WIDDIS, B.Sc.(Eng.), Associate Member.

The paper was first received 29th August, and in revised form 8th November, 1955. It was published in February, 1956, and was read before the NORTH-WESTERN MEASUREMENTS GROUP *20th March, and the* MEASUREMENT AND CONTROL SECTION *17th April, 1956.)*

SUMMARY

The paper is an investigation into the possibilities of using an ndirectly heated thermistor as a precise a.c./d.c. transfer device over wide range of frequencies. The thermal drift inevitable with such a levice has been examined and a procedure for its elimination developed, esulting in a very high degree of thermal stability. Experiments made n the reproducibility of the thermistor seem to indicate a complete bsence of hysteresis over the working range.

It is shown that the device has an extremely high sensitivity and is herefore useful for general laboratory work with simple equipment, lthough this advantage is somewhat outweighed by the slow response.

Experimental results are given for the d.c. reversal errors due to 'eltier and Thomson effects; these are found to be quite small.

An analysis of the effects of frequency shows that, owing to the rge thermal time-constant, the indirectly heated thermistor in its resent form can be used at frequencies as low as 0·2 c/s, thus well elow the working range of normal measuring devices. The response very slow owing to this large time-constant, and methods of construction are suggested which would improve the speed of response at ower frequencies and above.

LIST OF PRINCIPAL SYMBOLS

R = Resistance of the thermistor bead.
T = Absolute temperature, °K.
θ = Temperature, °C.
θ_c = Peak value of the temperature fluctuations, °C.
a, b = Thermistor parameters.
α_θ = Temperature coefficient of the thermistor bead at temperature θ°C.
K = Dissipation constant, watts/°C.
α_2 = Temperature coefficient of K.
d = Density, g/cm³.
s = Thermal capacity, joules/g/°C.
k = Thermal conductivity, watts/cm/°C.
P = Power dissipation, watts.
δP = Incremental power dissipation.
I = Heater current, amp.
δI = Incremental heater current.
$\delta\theta$ = Increment in bead temperature.
$\delta\theta'$ = Increment in ambient temperature.
R_h = Heater resistance.
α_0 = Temperature coefficient of the heater.
V = Bridge voltage.
α_B = Bridge temperature coefficient.
a_1 = Area of cross-section of heater wire, cm².
p = Instantaneous power.
m = Mass, grammes.
a_2 = Surface area of bead, cm².

Mr. Widdis is in the Department of Electrical Engineering, Northampton Polytechnic, ondon.

(1) INTRODUCTION

The rapid developments in electrical engineering practice necessitate an increasing degree of accuracy in the precise measurement of alternating voltages and currents over a wide frequency range. The basis of all such measurements is a transfer device which can be calibrated on direct current and has a known, and preferably small, error when used on alternating current. The actual a.c./d.c. transfer devices in general use are:

(*a*) The electrostatic voltmeter which has been used at the National Physical Laboratory for many years. Recent improvements in this type of voltmeter now enable it to be used at frequencies as high as 20 kc/s. The elaborate nature of the equipment and its accessories, however, render it unsuitable for use outside a national standardizing laboratory.

(*b*) The reflecting electrodynamic instrument which is used in many forms. This has, however, a number of disadvantages, the principal ones being the high power consumption and limited frequency range.

(*c*) The lamp bridge, which has been of little use owing to the difficulties in eliminating the harmonics which arise as a result of the non-linear nature of a filament lamp used as a circuit element.

(*d*) The thermal convertor utilizing a heated thermocouple, which has been in general use for a number of years; limitations arise, however, since these units are temperature dependent.

The errors appearing in thermal convertors such as vacuo-thermo-junctions have been exhaustively investigated by Hermach[1] of the National Bureau of Standards, Washington, D.C., and this work has led to their adoption at the N.B.S. as fundamental standardizing devices.

A thermal convertor in the form of a vacuo-thermo-junction appears to be an attractive a.c./d.c. transfer device for general use. There are, however, certain limitations inherent in these

devices which require further discussion. Reversal errors appear when used on direct current if the thermo-junction is not at the mid-point of the heater. Carefully conducted experiments on a number of commercial units show reversal differences ranging from 0·01% to 0·2%, and although the mean of the forward and reverse readings may be taken for a.c. calibration, a large difference causes experimental difficulties, 0·03% usually being the maximum difference tolerable. This necessitates selection of the units, and experience shows that about one in four commercial couples will have reversal differences of less than 0·03%. The Thomson effect causes asymmetric temperature distribution on direct current and gives a transfer error. This has been estimated by Hermach to be about 0·015% for a heater made from an alloy of 90% Ni and 10% Cr, and to be less than 0·005% for a manganin heater unit. In general, materials with high Thomson coefficients are unsuitable for use as heater materials. In addition to this, errors appear at low frequencies owing to the non-linear response of the couple if the thermal inertia of the heater is low. If a high degree of sensitivity is to be achieved, the mass of the heater should be small, and this restricts the choice of heater materials to those having a low thermal diffusivity λ, given by $\lambda = k/ds$. Again manganin appears to be the most suitable heater material.

The sensitivity of the vacuo-thermo-junction is low, the output being about 6 mV in a few ohms at full rated current. Hence, for a definition of 1 part in 10^4 to be achieved, the galvanometer must be capable of detecting 0·5 μV, and considerable care must be exercised to avoid stray thermal e.m.f.'s. These units are sensitive to changes in ambient temperature, but if supplies having a stable voltage are available, and measurements can be effected with rapidity, the drift is not serious. Hermach gives a figure of about 0·05% per hour for the drift rate in his tests, Carefully conducted tests in a normal laboratory on units of English manufacture gave a maximum drift rate of 0·003% per minute with the units protected from draughts, but without any attempt to control the temperature.

Thermal drift is undesirable and attempts have been made to eliminate it by using identical vacuo-thermo-junctions in a back-to-back arrangement, an oil bath having been adopted by Rump[2] to ensure equality of temperature.

The author's experience with the various devices referred to prompted this experimental investigation, whose primary object was the provision of a stable and sensitive a.c./d.c. transfer device for use with an audio-frequency a.c. potentiometer. It was noted during the course of the work that the indirectly heated thermistor is a useful measuring device at very low frequencies.

(2) THE INDIRECTLY HEATED THERMISTOR

This device has obvious possibilities as an a.c./d.c. transfer element if it will give reproducible results. Experience with the very high degree of stability achieved in the Patchett a.c. stabilizer,[3] which utilizes a thermistor bridge, suggested that the reproducibility of the thermistor would be good and prompted the initiation of an investigation into its possibilities. Schrader[4] has described a thermistor bridge used as an a.c./d.c. transfer device, but does not appear to have investigated its behaviour in any detail.

The indirectly heated thermistor, as shown in Fig. 1, comprises a bead of a special material having a relatively large negative temperature coefficient, mounted between platinum leads.[5] The bead is surrounded by an insulated 100-ohm heater of nickel-chrome alloy, and the whole is in a sealed evacuated glass envelope. Experiment shows that the bead material has a large negative temperature coefficient of resistance of as much as 0·04, and as a result small changes in the heater current cause large changes in the bead resistance. These changes in the bead

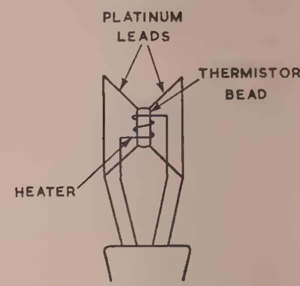


Fig. 1.—Schematic arrangement of indirectly heated thermistor.

resistance provide a very sensitive indication of heater curren changes. Clearly, if this device is to be used as an a.c./d.c transfer element, very good compensation against ambient tem perature changes is essential. A balanced bridge arrangemen using two thermistors is the obvious solution, but in practic complications occur owing to small differences in the temperatur coefficients of the thermistor beads, and difficulties in ensurin that both are subject to the same temperature changes.

Schrader found that a thermistor bridge constructed wit thermistors of Continental origin, still showed overall temperatur drift owing to these effects and necessitated compensation by th introduction of resistance material having a positive temperatu coefficient into one bridge arm.

However, experience with thermistors of English manufactu shows that, normally, a simple bridge arrangement is entire satisfactory, provided that certain conditions of operation a fulfilled.

(2.1) Method of Operation

The actual bridge arrangement used in this investigation illustrated in Fig. 2, T_A and T_B being two indirectly heate

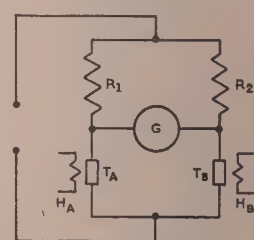


Fig. 2.—Bridge arrangement of two indirectly heated thermistors.
H_A, H_B = Heaters.
T_A, T_B = Thermistor beads.

thermistors of the same type. R_1 is a fixed resistor and R_2 is resistor variable over a limited range for initial balancing pu poses. The bridge supply is from a 9-volt dry battery and t detector is a portable reflecting galvanometer. The operati current range for the heaters is preferably between 2·5 and 15 m

When used as a standardizing device as in a.c. potentiometry known direct current is passed through the two heaters, H_A a H_B, in series and the bridge is balanced by adjustment of R Then H_A is switched to the a.c. circuit with the direct current H_B maintained at its correct value. The alternating current then adjusted until the bridge is again balanced. If the unit is fr from transfer error the r.m.s. value of the alternating current equal to the direct current. An alternative procedure is desiral if an alternating current has to be measured. In this case it better to balance the bridge first with alternating current in bo heaters, then change one heater to direct current which is adjust until balance is achieved and its magnitude then measur Accessory apparatus comprises means for regulating the dir

current, a standard resistance and a precision d.c. potentiometer for its measurement.

(3) THE EFFECT OF AMBIENT TEMPERATURE CHANGES

It has been shown experimentally by several investigators that within close limits the law relating the resistance R of a thermistor bead to its absolute temperature T is

$$R = a\varepsilon^{b/T} \quad \dots \dots \quad (1)$$

where a is a constructional constant and b is a constant determined by the bead composition. If the temperature coefficient of resistance at temperature $\theta°$ C is defined as

$$\alpha_\theta = \frac{1}{R}\frac{dR}{d\theta} \quad \dots \dots \quad (2)$$

it follows that the temperature coefficient is

$$\alpha_\theta = -\frac{b}{T^2} \text{ or } -\frac{b}{(273+\theta)^2} \quad \dots \quad (3)$$

Clearly, with beads having the same value for b, which is a reasonable expectation with units of the same type, it follows that when the two thermistor beads in a bridge are at the same temperature they must have equal temperature coefficients. Thus, if this condition could be satisfied the bridge would remain balanced no matter what changes occurred in the ambient temperature.

Owing to the inherent constructional difficulties it is inevitable that any pair of thermistors will differ both in their bead-circuit resistances and in their heater-circuit resistances, and it is necessary to achieve balance on the bridge while satisfying the condition of equal temperature as closely as possible. The procedure actually adopted kept the bridge current small so that it had only a small effect on the bead temperature, and then one of the heaters was shunted so that the same bead temperature was attained in the two thermistors for the same heater current. Complete information on the thermistor operating characteristics was necessary.

(3.1) Measurement of the Thermistor Characteristics

Series of measurements were carried out on pairs of thermistors which were not selected in any way. The general results were similar, so for simplicity the discussion is restricted to a single pair.

The bead resistances were measured at various temperatures by immersing the thermistors in a heated oil bath, and measuring the resistances on a Wheatstone bridge with a very small bridge current. A plot of log R against $1/T$ as shown in Fig. 3 gave a straight line, thus enabling the constants b and a to be determined. b is given by 2·303 times the gradient of the line and a is then calculated from eqn. (1). With b and a known, measurements of the bead resistance at ambient air temperature with various values of heater current flowing enabled the bead temperature to be calculated from eqn. (1) for the various values of power dissipated in the heater. The resulting plot of temperature against the power dissipation, in milliwatts, in the bead is

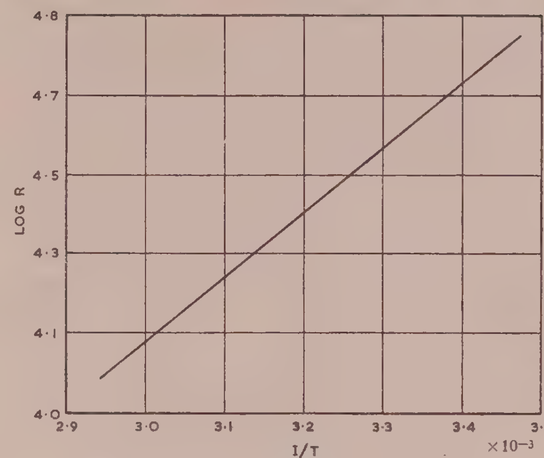


Fig. 3.—Relationship between log R and $1/T$.

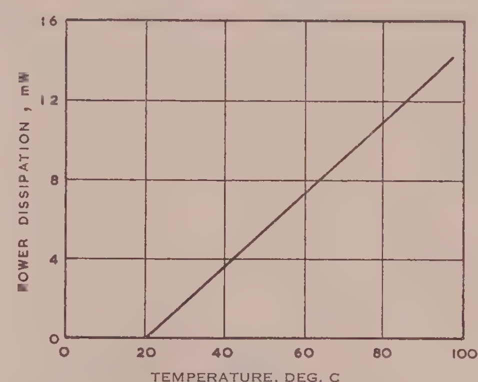


Fig. 4.—Temperature rise, in degrees centigrade, against bead power dissipation, in milliwatts.

shown in Fig. 4, and is practically linear up to a power dissipation of 10 mW—the maximum value adopted for working conditions. The dissipation coefficient, K, in milliwatts per degree centigrade, may therefore be taken as constant. The temperature coefficient of resistance, α_θ, can be calculated from these results by the use of eqn. (3). Typical results for the Type B.5412/60 thermistors are given in Table 1.

Table 1

TYPICAL CONSTANTS FOR THERMISTORS TYPE B.5412/60

Unit	R at 20·2° C	a	R_h	K	b	T at 10 mW dissipation	α_θ at 10 mW dissipation
	Ω	Ω	Ω	mW/° C	°K	°K	
A	55 680	0·122	99·4	0·181	3 820	348	0·031
B	57 310	0·125	100·5	0·183			

(3.2) Determination of the Optimum Working Condition

A plot of the bead circuit resistance against heater dissipation for two Type B.5412/60 thermistors is given in Fig. 5. It can be seen that the curves are very close together. It was found experimentally that the same resistance/energy dissipation results were obtained when the bead was heated by passing current through it, as when the heater was used.

Taking a dissipation of 10 mW for reference, the actual bridge operating point may be found by plotting the voltage across the bead against bead current. The values can be readily obtained from the resistance/dissipation curves since

$$\delta P \doteqdot V_1^2/R = I_B^2 R$$

where δP is the incremental dissipation above the heater power, P; V_1 is the voltage across the bead; I_B is the bead current; and R is the bead resistance corresponding to the total dissipation $P + \delta P$. The results for the two thermistors A and B are shown in Fig. 6. The applied bridge voltage is represented by OC, and a line CD is drawn where OC/OD $= R_1$. The point of intersection G of

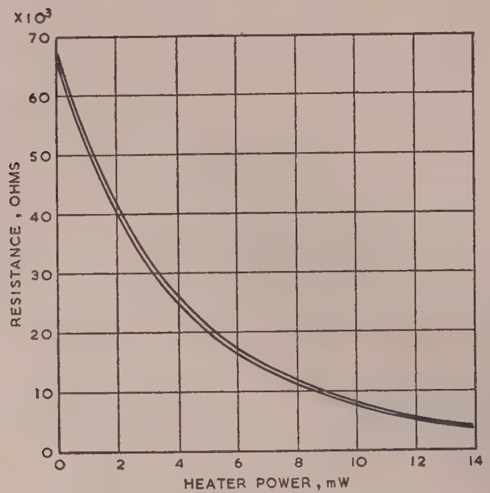


Fig. 5.—The relationship between bead resistance and heater power for two Type B 5412/60 thermistors.

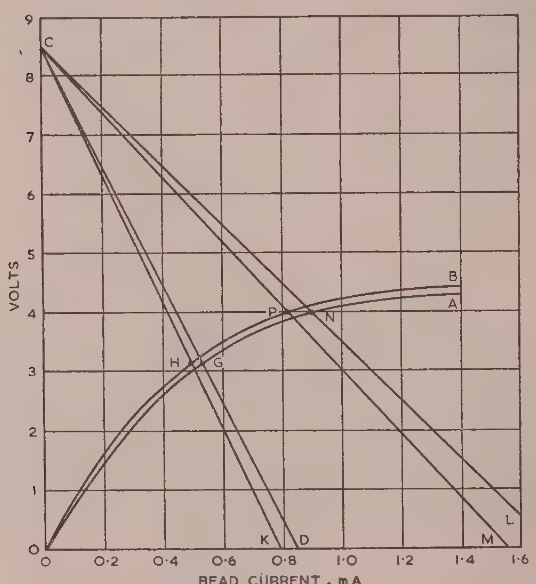


Fig. 6.—Bridge operating conditions for 10 mW initial dissipation. (Air temperature, 16·6° C.)

the line CD with curve A is the operating point for thermistor A. When the bridge is balanced the voltages across the thermistor beads must be equal, and the operating point H for the thermistor B is obtained by drawing GH parallel to the abscissae. The gradient of line CHK is then the resistance R_2. Clearly GH represents the difference in power dissipation in the two beads when the bridge is balanced.

In the particular conditions shown in Fig. 6, the bridge supply is 8·5 volts and line CGD is drawn for a resistance of 10 000 ohms. Line CNL is drawn for a resistance of 5 000 ohms, and the resistance in series with thermistor B is given by the slope of CPM. Again PN represents the difference in power dissipation in the two beads for this condition, and it is apparent that the difference in bead dissipation increases as the bridge arms are reduced in value.

A family of voltage/current curves for differing values of heater

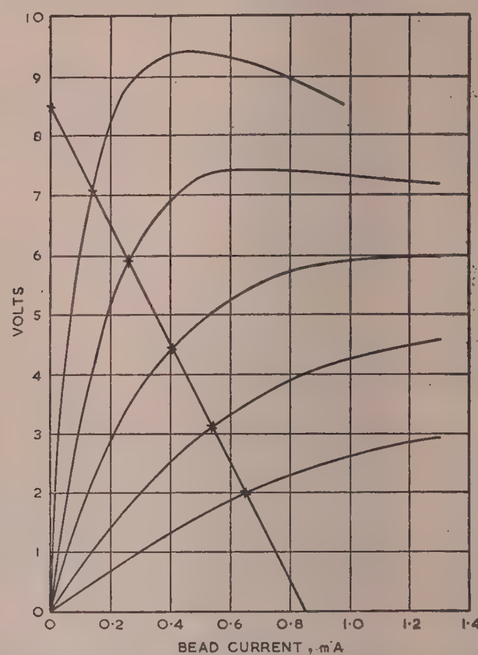


Fig. 7.—Thermistor characteristics for various values of heater curren
(Ambient temperature, 19·8° C.)

An operating line for a series resistance of 10 000 ohms and bridge supply of 8·5 vol is inserted.

current is given in Fig. 7. A working line has been inserted, an for stable operation of the unit at low heater currents the workin lines must lie within the hump in the characteristic.

(3.3) Derivation of the Bridge Temperature Coefficient

The bridge temperature coefficient may conveniently defined as the fractional change in heater current required restore balance per degree centigrade rise in ambient temperatur

$$\alpha_B = \frac{\delta I}{I}\frac{1}{\delta\theta'} \quad . \quad . \quad . \quad . \quad . \quad . \quad ($$

and an expression for this may be derived as follows.

The two thermistor beads will have slightly different operati temperatures owing to the unavoidable difference in bead dissip tion when the bridge is balanced. If one bead is at T°K a the other at $T + \delta\theta$°K the difference in their temperatu coefficients is

$$\delta\alpha = \frac{2\delta\theta}{T}\alpha \quad . \quad . \quad . \quad . \quad . \quad .$$

Now, if the temperature of the whole bridge is increased by the change in temperature of one bead, $\delta\theta''$, required to rest balance is $\delta\theta'' = \delta\theta' \times \delta\alpha/(\alpha - \delta\alpha)$ or to a close approximati

$$\delta\theta'' = \frac{\delta\alpha}{\alpha}\delta\theta' \quad . \quad . \quad . \quad . \quad . \quad .$$

Substituting from eqn. (5) in eqn. (6) we have

$$\delta\theta'' = \frac{2\delta\theta}{T}\delta\theta' \quad . \quad . \quad . \quad . \quad . \quad .$$

Since the power dissipation in the heater is $P = I^2R_h$ it foll that a small change in heater current δI produces a dissipat change

$$\delta P = 2I\delta IR_h \quad . \quad . \quad . \quad . \quad . \quad .$$

Now a change in heater current, δI, will change the temperature of one bead by $\delta\theta''$, so that

$$\delta\theta'' = \frac{\delta P}{K} = \frac{2I\delta IR_h}{K} \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

If this value of $\delta\theta''$ is used in eqn. (7), we get

$$\delta\theta' = \frac{I\delta IR_hT}{K\delta\theta} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

and if this is put in eqn. (4) we get

$$\alpha_B = \frac{K\delta\theta}{I^2R_hT} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

The difference $\delta\theta$ in bead temperatures is dependent upon the difference in dissipation constants, heater resistances and operating points of the two thermistors.

The contribution to α_B due to the difference in the operating points referred to in Section 3.2 can be assessed by supposing that $\delta\theta$ is due to the difference in bead dissipation, ΔP, at the operating point. Then

$$\delta\theta = \Delta P/K$$

giving

$$\alpha_{B1} = \Delta P/R_hI^2T \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

Substitution of numerical values in eqn. (11) gives a value for α_B of $0{\cdot}45 \times 10^{-4}$ per degree centigrade difference in bead temperature. A 2% difference in the values of K for the two thermistors gives a temperature difference of $1{\cdot}1°$ C at 10 mW dissipation, and it is apparent that substantial differences in the values of K can be tolerated.

Differences in the values of K and heater resistance for the two thermistors can be compensated by shunting one of the heater resistances. Let the two units have heater resistances R_{h1} and R_{h2} and dissipation constants K_1 and K_2 respectively. If $R_{h2}/K_2 > R_{h1}/K_1$, a condition of equal bead temperature for all heater currents can be obtained by shunting R_{h2} by a resistance r whose value is given by

$$r = R_{h2}\left(\frac{q + \sqrt{q}}{1 - q}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

where $q = R_{h1}K_2/R_{h2}K_1$

If this condition is met, α_B is given by eqn. (12), but the curve as in Fig. 6, for the shunted unit, should be replotted with an initial dissipation corresponding to the shunted condition before deriving ΔP.

The procedure outlined will, in general, result in the attainment of a very low value of α_B.

(3.4) Experiments on Ambient Temperature Effects

Since the actual temperature coefficient of each thermistor at the working temperature is about $0{\cdot}03$, it is essential for both thermistors to be affected by ambient temperature changes to exactly the same extent. This may be ensured by surrounding the two thermistors by a large thermal mass. Previous investigators have used petroleum baths for this purpose, but in this instance large variations were encountered in the balance when the bridge was immersed in such a bath. It would appear that immersion in a liquid can be successful only if convection currents in the liquid can be avoided. The somewhat simpler solution adopted was to insert the two thermistors in close-fitting holes drilled in a 1 lb brass block, the holes then being filled with oil. The brass block was surrounded by asbestos lagging to minimize the effects of draughts. This arrangement achieved a very high degree of stability, and it was found that in a normal laboratory, i.e. with the air temperature uncontrolled, the drift in the bridge balance amounted to less than $1\frac{1}{2}$ parts in 10^4 even over a period as long as 60 hours.

A bridge supply voltage of $8{\cdot}5$ volts was adopted. The bridge resistance in series with thermistor A was fixed at 10 000 ohms, and with a power dissipation of 10 mW in both heaters, balance was achieved with a resistance of 10 400 ohms in series with B.

A condition of equal power dissipation in the heaters was adopted in this test for equal currents in the heater circuits. The resistances of the two heaters differed by $1{\cdot}1$ ohm, and this condition was met by shunting the heater with the higher resistance (thermistor B), the value of the shunt being given by eqn. (13) with $q = R_{h1}/R_{h2}$.

Measurement of the change in balance current with ambient temperature changes gave a value of $0{\cdot}43 \times 10^{-4}$ per degree centigrade for α_B. The difference in bead dissipation at the operation point obtained from the curves was $0{\cdot}07$ mW, and with a bead operating temperature of 357° K, eqn. (12) gives a contribution to α_B, owing to this, of $0{\cdot}17 \times 10^{-4}$ per degree centigrade. The additional contribution due to the differences in the measured values of K for the two beads was calculated to be $0{\cdot}3 \times 10^{-4}$ per degree centigrade giving a total calculated value for α_B of $0{\cdot}47 \times 10^{-4}$ per degree centigrade.

This is a reasonable agreement and indicates that the theoretical conditions for equal ambient temperature change can be met.

It so happened that with this particular pair of thermistors the characteristics were such that a lower value of α_B could be achieved without the shunt, but this may be regarded as fortuitous.

(4) BRIDGE SENSITIVITY

An expression for the bridge detector current can be derived simply by considering the bridge to have equal ratios, when, for

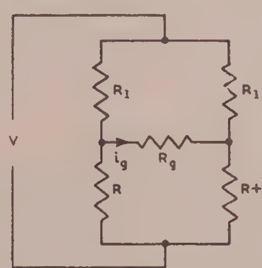


Fig. 8.—Network for sensitivity calculation.

the arrangement in Fig. 8, the detector current i_g for a small change of resistance in one arm from R to $R + \delta R$ is

$$i_g = \frac{VR_1\delta R}{R_g(R_1+R)^2+2RR_1(R_1+R)+\delta R[R_1^2+2RR_1+R_g(R_1+R)]} \quad . \quad . \quad . \quad . \quad (14)$$

Now, by definition

$$\alpha_\theta = \frac{1}{R}\frac{dR}{d\theta}$$

i.e.

$$\delta R = \alpha_\theta R\delta\theta$$

where

$$\delta\theta = \delta P/K = 2I\delta IR_h/K$$

giving

$$\delta R = \frac{\alpha_\theta R2I\delta IR_h}{K} \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

substituting in eqn. (14) and neglecting the terms in δR in the denominator gives

$$i_g = \frac{VR_1R}{R_g(R_1 + R)^2 + 2RR_1(R_1 + R)} \times \frac{\alpha_\theta 2I\delta IR_h}{K} \quad . \quad (16)$$

α_θ varies with the operating current, but for the calculation of sensitivity, which need only be approximate, it may be taken as $0{\cdot}03$.

If, with 9 volts applied to the bridge, a 450-ohm 2 sec portable galvanometer with a sensitivity of 250 mm/μA is used, and the values inserted in eqn. (16) are as follows

$$\left.\begin{aligned} R_1 &= 10\,000 \\ R &= 6000 \\ R_h &= 100 \end{aligned}\right\} \text{ohms.} \qquad \begin{aligned} \alpha_\theta &= 0{\cdot}03 \\ K &= 0{\cdot}2 \times 10^{-3}\,\mathrm{mW/^\circ C} \\ I &= 10 \times 10^{-3}\,\mathrm{amp} \end{aligned}$$

we find that for $\delta I/I = 1 \times 10^{-4}$, $i_g = 0{\cdot}08\,\mu$A, giving a scale deflection of 20 mm.

A similar computation for an operating current of 5 mA gives a deflection of 2·5 mm for $\delta I/I = 1 \times 10^{-4}$.

It is apparent that the sensitivity of the device is extremely high.

It is interesting to compare the sensitivity of a vacuum-thermo-junction. In this case the output voltage v is proportional to I^2, written as $v = CI^2$, where C is a constant and

$$\delta v = C \, . \, 2I\delta I$$

For a typical unit, v would be 6 mV at full rated current of 10 mA, giving $\delta v = 1{\cdot}2\,\mu$V when $\delta I/I = 1 \times 10^{-4}$. The couple resistance is 9 ohms, and in this case a 10-ohm 2 sec portable galvanometer, having a sensitivity of 23 mm/μA would be suitable, giving a scale deflection of 1·4 mm.

Hence for 10 mA operation the thermistor bridge is about 14 times as sensitive as a vacuo-thermo-junction.

(5) REPRODUCIBILITY

Serious difficulties can arise in thermal transfer devices if hysteresis is present, i.e. if the characteristics of the device change after having been taken through a temperature cycle. A limited amount of information on ageing has been produced by Pearson,[5,6] who gives a figure of 0·2% change in resistance after one year for an unenclosed unit and mentions that it is less for enclosed units.

A series of experiments was carried out on two types of thermistor to investigate their reproducibility. The thermistors used for this test had nominal bead resistances of 50 000 ohms and 500 ohms at 20° C. Pairs of indirectly heated thermistors were arranged in the bridge circuit of Fig. 2, and the same direct current was passed through both heaters, each being supplied from a separate circuit. Changes in the bridge balance which occurred after interrupting one heater circuit for varying periods up to 2 min were noted. No change in balance which could be attributed to ageing was noticed. Over a two-hour period the maximum deviation in balance was equivalent to heater-current changes of ± 5 parts in 10^5, which was about the limit of experimental error. It can be inferred, however, that for a given unit, a reproducibility of rather better than 1 part in 10^4 is readily attainable for heater dissipations up to 10 mW. Although a similar order of reproducibility was achieved for heater dissipations up to 40 mW, it was felt that 20 mW was the maximum figure desirable.

These reproducibility experiments appeared to indicate that the low-resistance thermistors were not quite as consistent as the high-resistance ones, hysteresis effects being noticeable when the heater dissipation exceeded 10 mW.

It should be noted that the thermistors were not subjected to any special ageing process before these tests were performed, but were used as received from the manufacturer.

(6) A.C.–D.C. TRANSFER ERRORS

Any device which is intended to be used as a transfer element should have the same response for r.m.s. alternating current as it has for direct current. It would appear that a thermal device such as a vacuo-thermo-junction, or an indirectly heated thermistor of the type under investigation, should satisfy this condition fully. In practice, however, transfer errors appear for various reasons. These are Peltier and Thomson effects appearing in the d.c. test, and in the a.c. test integration errors arise owing to temperature cycles in a non-linear circuit element and errors due to the shunt capacitance of the heater.

(6.1) Errors due to Peltier and Thomson Effects

When a direct current is passed through a length of fine resistance wire mounted between two supports, the resulting temperature distribution along the wire is asymmetrical (*a*) because of Peltier effects at the junction of the wire with the supports, and (*b*) because of Thomson heating along the wire. These effects cause a difference in the response to direct current depending upon the direction of flow, the so-called reversal error, but will not appear during the alternating-current test.

These effects can cause significant errors in a device like the vacuo-thermo-junction, which consists of a length of heated wire between supports with a thermocouple measuring the mid-point temperature—this being between 150 and 200° C—and they have been discussed in the Introduction.

The indirectly heated thermistor can be operated at a much lower temperature, 70° C above ambient temperature having been adopted as a maximum for these tests, and as a result these effects are not so significant. The greater part of the heater wire is in contact with the bead, and the problem of finding the exact mid-point of the heater during construction does not arise. As the supports are relatively massive the Peltier effect will generally cause a negligible change in the support temperature and can be ignored.

The Thomson coefficient, in volts per degree centigrade, is given by the derivative of the thermo-electric power equation for the heater material against lead, multiplied by the absolute temperature. The short length of heater material between the support and the bead is subject to a temperature rise of nearly 70° C. Approximate estimates of the Thomson heating, assuming the Thomson effect to be constant, indicate that it may amount to 6 μW, which will be additive in one lead and subtractive in the other. Clearly, as the total heater power dissipation is $10^4\,\mu$W, this amount is significant and there will be a small temperature difference between the ends of the bead owing to this, and if there is constructional asymmetry it will probably cause the reversal errors encountered.

All the thermistors examined had heaters of nickel–chromium which has a rather high Thomson coefficient, but the reversal errors encountered are extremely small and it is reasonable to assume that these represent the maximum transfer errors due to this cause. This assumption was confirmed by a comparison

Table 2

FRACTIONAL DIFFERENCES BETWEEN FORWARD AND REVERSE DIRECTIONS OF D.C. HEATER CURRENT FOR THE SAME BEAD RESISTANCE AT 10 MW HEATER DISSIPATION

Unit	A	B	C	D	E	G	H
Reversal difference, $\delta I/I$	$0{\cdot}5 \times 10^{-4}$	$1{\cdot}0 \times 10^{-4}$	$1{\cdot}7 \times 10^{-4}$	$1{\cdot}0 \times 10^{-4}$	$1{\cdot}25 \times 10^{-4}$	$0{\cdot}2 \times 10^{-4}$	$0{\cdot}7 \times 10^{-4}$

est with a selected vacuo-thermo-junction discussed later in the aper. The reversal differences for seven units are given in able 2. Some improvement could be achieved by the use of nanganin as a heater material, but it is felt that this refinement vould be hardly justifiable in view of the results obtained.

6.2) Errors appearing in A.C. Tests owing to Non-Linearities

A thermal device such as the indirectly heated thermistor will ntegrate correctly on alternating current only if the mean bead esistance attained is the same as that in the d.c. test. The relationhip between bead resistance and temperature is non-linear, with he result that, if periodic fluctuations occur in the bead temerature, there will be a change in the mean resistance. At igh frequencies the thermal inertia of the bead will prevent any emperature fluctuations, but at low frequencies, such fluctuations ppear and can cause an additional transfer error. In addition, he heat loss from the bead varies as the fourth power of temerature, and temperature fluctuations will cause an additional eat loss, further altering the mean resistance.

6.2.1) Change in Mean Resistance Due to Temperature Oscillations in a Non-Linear Element.

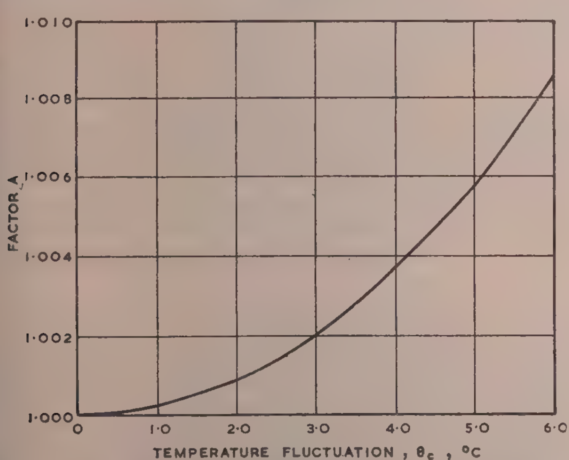
Assuming sinusoidal temperature fluctuations, an expression as been derived in Section 12.1 for the mean resistance on lternating current when the resistance is subjected to temperature uctuations, giving the mean resistance

$$R = a\varepsilon^{\frac{b}{T}}\varepsilon^{\frac{b\theta_c^2}{2T^3}}\left[I_0\left(\frac{b\theta_c}{T^2}\right)I_0\left(\frac{b\theta_c^2}{2T^3}\right)\right] \quad . \quad . \quad . \quad (17)$$

vhere T° K is the mean temperature corresponding to the d.c. ondition.

The factor $$A = \varepsilon^{\frac{b\theta_c^2}{2T^3}}\left[I_0\left(\frac{b\theta_c}{T^2}\right)I_0\left(\frac{b\theta_c^2}{2T^3}\right)\right] \quad . \quad . \quad . \quad (18)$$

s a measure of the deviation of the mean temperature between he a.c. and d.c. tests in so far as A departs from unity. Values f A have been calculated and are plotted in Fig. 9 for a peak emperature fluctuation of $\pm 0{\cdot}5$ to $\pm 10^\circ$ C.



ig. 9.—Relationship between factor A and the peak value of the temperature fluctuations.

This treatment assumes that the whole mass of the bead is ubject to the same fluctuations; this is not strictly valid, but he expression gives a reasonable indication of the magnitude of rror likely to exist from this cause at low-frequency operation. The effect is more usefully expressed in terms of fractional change in heater current to restore balance, and this may be deduced as follows:

Put $$R_{ac} = R\left(1 + \frac{\delta R}{R}\right) = RA \quad . \quad . \quad . \quad . \quad (19)$$

so that $$\frac{\delta R}{R} = A - 1$$

when, by definition, $$\frac{\delta R}{R} = \alpha_\theta \delta\theta = \alpha_\theta \frac{\delta P}{K}$$

but $$\delta P/P = 2\delta I/I$$

therefore $$\delta I/I = \delta P/2P = \frac{K\delta R}{\alpha_\theta R 2P}$$

giving $$\delta I/I = \frac{K(A-1)}{2P\alpha_\theta} \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

Values of $\delta I/I$ are given in Table 3.

An expression is developed later in the paper for the magnitude of the temperature oscillations which occur in the bead when the heater is energized with alternating current. The magnitude of the fluctuations can also be experimentally assessed by energizing the heater with alternating current, passing a small direct current through the bead and measuring the voltage fluctuations across it with the aid of a sensitive cathode-ray oscillograph; this enables the resistance fluctuations to be calculated, and these can be converted to equivalent temperature fluctuations with the aid of eqn. (2). An actual test at 1 c/s carried out on the two selected thermistors gave an equivalent temperature swing of $\pm 0{\cdot}37^\circ$ C.

(6.2.2) Additional Radiation Loss Due to Temperature Oscillations.

The heat loss by radiation from a surface is given by $\eta E(T^4 - T_0^4)$, where E is the surface emissivity, and if there are periodic fluctuations in T there will be an excess of radiation loss over the steady value. It is shown in Section 12.2 that the increase in power loss due to this effect is given by

$$\delta P = 12{\cdot}1 \times 10^{-12} a_2 T^2 \theta_c^2 \text{ watts} \quad . \quad . \quad . \quad (21)$$

This gives

$$\delta I/I = \delta P/2P = \frac{6{\cdot}05 \times 10^{-12} a_2 T^2 \theta_c^2}{P} \quad . \quad . \quad (22)$$

The area of bead surface, a_2, is approximately $0{\cdot}1\ \text{cm}^2$, and taking P to be 10 mW, we find that if $\delta I/I$ is not to exceed one part in 10^4, θ_c must be less than $3{\cdot}65^\circ$ C.

(6.2.3) The Relationship between Temperature Fluctuations and Supply Frequency.

The instantaneous power supplied to the bead must equal the rate of storage of energy plus the rate of loss of energy by conduction along the wires and by radiation from the surface. Since, however, the heat loss by conduction is predominant, it follows that, to a first approximation, radiation loss can be neglected. We therefore have

$$p = \hat{I}^2 R_h \sin^2 \omega t = ms\frac{d\theta}{dt} + K(\theta - \theta_0) \quad . \quad (23)$$

The solution of this equation is

$$\theta = \theta_0 + A_1\varepsilon^{-\frac{Kt}{ms}} + \frac{I^2R_h}{K} - \frac{I^2R_h \sin(2\omega t + \phi)}{\sqrt{(K^2 + 4\omega^2 m^2 S^2)}} \quad . \quad (24)$$

Under steady-state conditions the amplitude of the temperature oscillations is

$$\theta_c = \frac{I^2R_h}{\sqrt{(K^2 + 4\omega^2 m^2 s^2)}} \quad . \quad . \quad . \quad . \quad (25)$$

Since the ratio ms/K was found to be 9 in this case, K^2 can be neglected in comparison with $4\omega^2 m^2 s^2$; hence

$$\theta_c = \frac{I^2R_h}{2\omega ms} \text{ (if } \omega > 0{\cdot}2) \quad . \quad . \quad . \quad . \quad (26)$$

i.e. the amplitude of the temperature oscillations varies as the reciprocal of the supply frequency. As K is known, ms can be simply determined by measuring the thermal time-constant ms/K. This was found to be 9 sec for the Type B 5412/60 thermistor. The calculated value of θ_c from these results at a frequency of 1 c/s was 0·46° C compared with a measured value of 0·37° C.

The calculated value of θ_c should be higher than the measured value since the behaviour of the thermistor is not strictly represented by eqn. (23). In actuality the unit comprises a thermistor bead with platinum connecting wires, surrounded with alumina, over which the heater is wound. A complete mathematical treatment would necessitate the solution of a rather difficult heat-transfer problem. The problem can more easily be considered by using an electrical analogy.

The electrical analogy to eqn. (23) is given in Fig. 10(a). The capacitance C corresponds to the thermal capacity, $H = ms$,

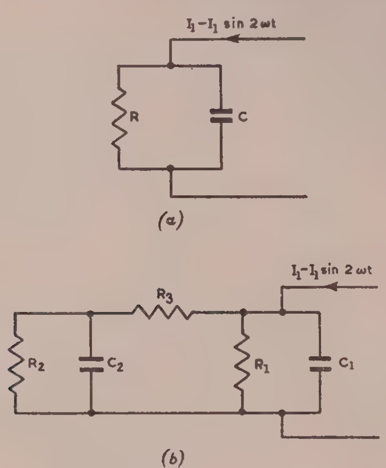


Fig. 10.—Electrical analogies to the heat-transfer problem.

and the resistance R corresponds to $1/K$. The circuit is supplied from a constant-current source, $I_1 - I_1 \sin 2\omega t$, where I_1 corresponds to I^2R_h and the voltage across the CR combination corresponds to the bead temperature. A closer analogy is given by the circuit shown in Fig. 10(b), also from a constant-current source. In this electrical analogy, C_1 corresponds to the thermal capacity of the heater and R_1 to the thermal resistance of the heater leads. R_3 corresponds to the thermal resistance of the alumina layer, and C_2 and R_2, respectively, represent the thermal capacity of the thermistor bead and the thermal resistance of its platinum leads. The voltage fluctuation V_m on C_2 represents the temperature fluctuations in the bead:

$$V_m = \frac{I_1R_1R_2}{\sqrt{\{(R_1+R_2+R_3-\omega^2C_1R_1C_2R_2R_3)^2+\omega^2[C_1R_1R_2+C_1R_1R_3+C_2R_1R_2+C_2R_2R_3]^2\}}} \quad . \quad . \quad . \quad (27)$$

While the use of this relationship gives a better correlation with the measured fluctuation, R_3 is fairly small, and it is considered that eqn. (26) is sufficiently accurate.

(6.2.4) **Additional Losses due to the Temperature Coefficient of Thermal Conductivity of the Leads.**

The dissipation constant K increases with temperature owing to radiation loss and to increase in the thermal conductivity of the leads. The effect of radiation loss has been considered, so in this case the dissipation constant can be written as $K = K_0(1+\alpha_2\theta)$.

The instantaneous bead temperature is $\theta = \theta_c \sin \omega t + \theta_M$.

Hence the instantaneous power dissipation is

$$\begin{aligned} p &= K_0(1+\alpha_2\theta)(\theta_c \sin\omega t + \theta_M) \\ &= K_0[1+\alpha_2(\theta_c\sin\omega t+\theta_M)][\theta_c\sin\omega t+\theta_M] \\ &= K_0(\theta_c\sin\omega t+\theta_M)+K_0\alpha_2[\theta_c\sin\omega t+\theta_M]^2 \quad (28) \end{aligned}$$

The average value of this is

$$P_{a.c.} = K_0\left(\theta_M+\alpha_2\theta_M^2+\frac{\alpha_2\theta_c^2}{2}\right) \quad . \quad . \quad . \quad (29)$$

On direct current

$$P_{d.c.} = K_0(\theta_M+\alpha_2\theta_M^2) \quad . \quad . \quad . \quad . \quad (30)$$

There is therefore an additional loss due to this factor on alternating current given by $K_0\alpha_2\theta_c^2/2$. The fractional loss, then, is

$$\frac{\delta P}{P} = \frac{K_0\alpha_2\theta_c^2}{2K_0(\theta_M+\alpha_2\theta_M^2)} = \frac{\alpha_2\theta_c^2}{2(\theta_M+\alpha_2\theta_M^2)} \quad . \quad . \quad (31)$$

This is equivalent to a fractional heater-current change

$$\delta I/I = \frac{\alpha_2\theta_c^2}{4(\theta_M+\alpha_2\theta_M^2)} \simeq \frac{\alpha_2\theta_c^2}{4\theta_M} \quad . \quad . \quad . \quad (32)$$

The supports to the thermistor bead consist of two heater wires of 80/20 nickel–chromium alloy, each of 0·001 2 in diameter, and four platinum leads each of 0·002 in diameter.

Now, the thermal conductance of a lead of length l and area of cross-section a_1, having one end at temperature θ_1 and the other at θ_0, is

$$\frac{a_1k}{l}\left[1+\frac{\beta}{2}(\theta_1+\theta_0)\right]$$

where β is the temperature coefficient of thermal conductivity. This relationship, in conjunction with the known values of k and β for platinum and Nichrome and the given dimensions, permits the estimation of a value for α_2 of 6×10^{-4}. The calculated values of $\delta I/I$ are given in Table 3. The value estimated for α_2 is probably within ±20%, but it can be seen that the error due to the presence of α_2 is so small that this accuracy is sufficient.

(6.2.5) **Effect of the Temperature Coefficient of Resistivity of the Heater.**

The resistance of the heater varies with temperature in accordance with the relationship $R_h = R_0(1+\alpha_0\theta)$. When the heater is supplied from an alternating-current source, the bead temperature is given by eqn. (24). If all the heater material is at the bead temperature,

$$R_h = R_0\left\{1+\alpha_0\left[\frac{\hat{I}^2R_0}{2K}-\frac{\hat{I}^2R_0}{2}\frac{\sin(2\omega t+\phi)}{\sqrt{(4\omega^2m^2s^2+K^2)}}\right]\right\} \quad (33)$$

Since α_0 is very small, sufficient accuracy is obtained by writing R_0 for R_h in the term in the inner brackets. When current $\hat{I}\sin\omega t$ is supplied to the heater, the instantaneous power is $p = \hat{I}^2R_h\sin^2\omega t$. Substituting for R_h, expanding and deriving the average power supplied, we get

$$P_{a.c.} = I^2R_0\left[1+\frac{\alpha_0I^2R_0}{K}+\frac{I^2R_0\alpha_0\sin\phi}{2\sqrt{(4\omega^2m^2s^2+K^2)}}\right] \quad . \quad (34)$$

On direct current

$$P_{d.c.} = I^2R_0(1+\alpha_0\theta) = I^2R_0\left(1+\alpha_0\frac{I^2R_0}{K}\right) \quad . \quad (35)$$

The fractional difference in power supplied is

$$\frac{P_{a.c.} - P_{d.c.}}{P_{d.c.}} = \frac{\delta P}{P} = \frac{+ I^2 R_0 \alpha_0 \sin \phi}{2\sqrt{(4\omega^2 m^2 s^2 + K^2)}} \quad . \quad . \quad (36)$$

Now

$$\sin \phi = \frac{K}{\sqrt{(4\omega^2 m^2 s^2 + K^2)}}$$

ence

$$\frac{\delta P}{P} = \frac{+ I^2 R_0 \alpha_0 K}{2(4\omega^2 m^2 s^2 + K^2)}$$

nd

$$\delta I/I = \frac{I^2 R_0 \alpha_0 K}{4(4\omega^2 m^2 s^2 + K^2)} \quad . \quad . \quad . \quad . \quad (37)$$

nserting numerical values and taking α_0 for Nichrome as 4×10^{-4} ives values for $\delta I/I$ of 4×10^{-7} at 1 c/s and 4×10^{-5} at 0·1 c/s. he error from this cause is therefore negligible compared with he previous errors considered.

.2.6) Summary of A.C. Effects Due to Non-Linearities.

The fractional increases in heater current $\delta I/I$ required to aintain balance on alternating current in the presence of the ffects considered in Section 6.2.1–6.2.5 have been calculated and e given in Table 3. It is apparent from these results that the on-linear resistance/temperature law in the thermistor bead is e predominant source of error, and the other errors can be eglected. It can be seen also that, if 0·1% is taken as the limit permissible error, the lower frequency limit is about 0·2 c/s. The positive sign indicates that the alternating current is eater than its correct value at balance.

(6.3) Electrical Constants of the Heater

The measured time-constant L/R of the heater was 4×10^{-9} sec nd appeared to be inductive. Approximate calculations of the lf-inductance of the heater gave a figure of 0·4 μH, so it can inferred that the self-capacitance of the winding is extremely w. It was not possible to measure the self-capacitance in the esence of this inductance and the comparatively low heater sistance, but it would be reasonable to assume from the easured time-constant that it would be far less than 20 $\mu\mu$F. he actual capacitance between the heater and bead was 3 $\mu\mu$F. The equivalent circuit of the heater is given in Fig. 11, the rrent in the R and L arm is

$$I_{RL} = \frac{V_2}{\sqrt{(R^2 + \omega^2 L^2)}}$$

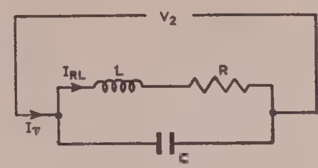


Fig. 11.—Equivalent circuit of heater.

and the total current in the circuit is given by

$$I_T = \frac{V_2\sqrt{[1 + \omega^2 C(CR^2 + \omega^2 L^2 C - 2L)]}}{\sqrt{(R^2 + \omega^2 L^2)}} \quad . \quad (38)$$

The departure of the term $\sqrt{[1 + \omega^2 C(CR^2 + \omega^2 L^2 C - 2L)]}$ from unity is a measure of the high-frequency error. Taking C as 20 $\mu\mu$F and L as 0·5 μH, we find that the departure from unity amounts to 1 part in 10^4 at 500 kc/s and 1 part in 10^3 at 1·5 Mc/s. There will also be a change in the effective resistance of the heater owing to skin effect. The heater-wire diameter is 0·001 2 in, and published figures for skin effect in single conductors show that a 0·002 in diameter nickel-chromium conductor changes its resistance by 1 part in 10^3 at 100 Mc/s. Skin effect can therefore be neglected in comparison with shunt-capacitance effects.

Table 3

Calculated A.C./D.C. Transfer Differences at a Heater Dissipation of 10 mW

Frequency	θ_c calculated	A.C./D.C. transfer error				
		1	2	3	4(*a*)	4(*b*)
c/s	°C					
0·1	4·6	$+16 \times 10^{-4}$	$+1·6 \times 10^{-4}$	$+0·5 \times 10^{-4}$	$-0·1 \times 10^{-4}$	$-0·43 \times 10^{-4}$
0·25	1·85	+2·6	+0·25	+0·08	−0·01	−0·07
0·5	0·92	+0·8	+0·06	+0·02	0	−0·01
1·0	0·46	0	+0·01	0	0	0

(1) Effect of non-linear law of resistance/temperature characteristic in bead.
(2) Effect of radiation.
(3) Effect of the temperature coefficient of thermal conductivity.
(4) Effect of the temperature coefficient of heater resistivity.
(*a*) $\alpha = 1 \times 10^{-4}$. (*b*) $\alpha = 4·3 \times 10^{-4}$.

(7) RESPONSE TIME

The response of the thermistor bridge is rather slow, and by the use of the transient term in eqn. (24) it is found that, starting with the thermistor cold, the time required to achieve balance to an equivalent heater-current difference of 1 part in 10^4 is theoretically 80 sec. In practice this time was found to be considerably more, the initial response time being modified by the attainment of thermal equilibrium with the surrounding mass. Fortunately, the time required for any measurement depends upon the initial condition of the thermistor, and if the "off" periods between operations are kept short, the slow response is not so noticeable.

When used, the bridge was set up initially and left for an hour before measurements were carried out. This was to permit both the bridge and the supplies to reach stability. It was found that the time required to carry out any test involving a heater switching operation was about 1 min for a precision of 1 part in 10^4, provided that switching was carried out rapidly. This time is certainly not excessive for measurements of this precision.

In the majority of applications, low-frequency response is unimportant and a considerable improvement in response time could be obtained by restricting the lowest operating frequency and effecting some changes in the thermistor design. Eqn. (20) gives a maximum value for θ_c of $1 \cdot 0°$ C for an error of 1 part in 10^4, and this can be used in conjunction with eqn. (24) for design purposes. The time-constant is given by ms/K and can be reduced by increasing K. This could be effected by reducing the length of the support leads through which most of the heat conduction takes place.

An increase in K would necessitate an increase in the heater operating current, which is desirable in the case of a.c. potentiometry where an operating current of the order of 50 mA is usual to keep the potentiometer resistance low.

A reduction in ms/K by a factor of 10 would greatly improve the response and would increase the lower frequency limit to 3 c/s. This, however, is a subject for future development.

It must be emphasized that very stable a.c. and d.c. supplies are essential for precise work; drift in the supplies is very serious and difficult to contend with in a device having a slow response.

In the comparison tests referred to in Section 8, a Patchett a.c. stabilizer and a very stable Wien bridge oscillator were used.

(8) COMPARISON TESTS

A series of comparison tests was carried out in which the same alternating current at frequencies between 25 c/s and 2 500 c/s was measured with the thermistor bridge and with a selected vacuo-thermo-junction. Agreement was obtained between the two devices to rather better than 2 parts in 10^4. The heater of the vacuo-thermo-junction was of nickel-chromium alloy, and the estimated error due to Thomson effect referred to in the Introduction was about 1 part in 10^4; when this was taken into consideration the a.c./d.c. transfer error for the bridge appeared to be $+ 1 \cdot 0$ part in 10^4.

The uncertainty in the results was estimated to be within ± 1 part in 10^4, so that it can be inferred that the transfer error is within $\pm 2 \cdot 5$ parts in 10^4. Considerable experimental difficulties arise in such comparison measurements owing to the limitations set by the stability of the supplies available.

(9) CONCLUSIONS

An investigation has been carried out into the performance of the indirectly heated thermistor as an a.c./d.c. transfer device, and a technique has been developed for the minimization of its inherent sensitivity to ambient temperature changes. Errors from all known causes have been assessed, and the results indicate that the performance compares favourably with alternative transfer devices. The exceptionally high sensitivity achieved means that a relatively insensitive portable galvanometer can be used as the detector, and this, coupled with the robust nature of the unit, makes it a very useful device for general laboratory work.

The indirectly heated thermistor, in its existing form, has a useful low-frequency performance. This, however, is linked with the rather slow response, which may be troublesome at higher frequencies if the supplies are not very stable, and it is felt that, for measurements in the power-frequency range and above, the constructional modifications suggested in the paper would result in an improved response with some loss in low-frequency range.

It is suggested that an error well within $\pm 2\frac{1}{2}$ parts in 10^4 may be expected over a very wide frequency range.

(10) ACKNOWLEDGMENTS

The work described in the paper was carried out as part of the research programme at the Northampton Polytechnic.

The author wishes to acknowledge the assistance received fron Dr. P. F. Soper, Head of the Electrical Engineering Department and Dr. A. W. Gillies.

(11) BIBLIOGRAPHY

(1) HERMACH, F. L.: "Thermal Convertors as A.C./D.C Transfer Standards for Current and Voltage Measurement at Audio Frequencies," *Journal of Research of the Nationa Bureau of Standards*, 1952, **48**, p. 121.

(2) RUMP, W.: "Über die genaue Absolutmessung von Wechsel-spannungen und einen Kompensator zur Prüfung vor Wechselstrom-Feinmessgeräten," *Elektrotechnik*, 1951 p. 64.

(3) PATCHETT, G. N.: "A New Precision A.C. Voltage Stabilizer," *Proceedings I.E.E.*, Paper No. 929 M, August, 1950 (97 Part II, p. 529).

(4) SCHRADER, H. J.: "Wechselstrom-Kompensationsmessunger mit indirekt geheizten NTC-Widerständen," *Elektro-technische Zeitschrift*, 1952, **A73**, p. 547.

(5) BECKER, J. A., GREEN, C. B., and PEARSON, G. L.: "Propertie and Uses of Thermistors—Thermally Sensitive Resistors," *Bell System Technical Journal*, 1947, **26**, p. 170.

(6) PEARSON, G. L.: "Thermistors—Their Characteristics anc Uses," *Bell Laboratory Record*, 1940–41, **19**, p. 106.

(7) GRAY, A., and MATHEWS, T. M.: "Treatise on Bessel Func-tions and their Applications to Physics" (Macmillan anc Co., Ltd., 1922), p. 36.

(12) APPENDICES—ANALYSIS OF THE LOW-FREQUENC ERRORS IN INDIRECTLY HEATED THERMISTORS

(12.1) Errors Due to the Non-Linear Relationship between Bea Resistance and Temperature

Since the resistance/temperature characteristic for the ther-mistor is non-linear, it follows that when the bead temperatur varies, the mean resistance may differ from its value at a mea temperature $T°$ K, obtained with a direct-current measurement.

Let the bead temperature have a small variation $\theta_c \sin \omega$ (where the amplitude θ_c is much less than T), whereupon th resistance of the bead is

$$R = a \exp \left[b \Big/ \left(1 + \frac{\theta_c}{T} \sin \omega t\right) T \right] \quad . \quad . \quad . \quad (39$$

If eqn. (39) is expanded by the binomial theorem and term above the second degree in θ_c are ignored, it will be found tha

$$R = a \exp \left[\frac{b}{T} \left(1 - \frac{\theta_c}{T} \sin \omega t + \frac{\theta_c^2}{T^2} \sin^2 \omega t + \ldots \right) \right] \quad . \quad (40$$

Now Sonine's expansion[7] gives

$$\exp (z \cos \theta) = I_0(z) + 2 \sum_{n=1}^{\infty} I_n(z) \cos n\theta \quad . \quad . \quad (4$$

where $I_n(z)$ is the modified Bessel function of the first kin If θ is replaced by $(\pi/2 - \theta)$,

$$\exp (z \sin \theta) = I_0(z) + 2I_1(z) \sin \theta$$
$$- 2I_2(z) \cos 2\theta - 2I_3(z) \sin 3\theta + \ldots \quad . \quad (4$$

Similarly, if θ is put equal to $\theta + \pi$,

since $\exp (-z \sin \theta) = \exp [z \sin (\theta + \pi)]$

$$\exp (-z \sin \theta) = I_0(z) - 2I_1(z) \sin \theta$$
$$- 2I_2(z) \cos 2\theta + \ldots \quad . \quad (4$$

Also, $\exp(z_1 \sin^2\theta) = \exp[\frac{1}{2}z_1(1-\cos 2\theta)]$

$$= \exp\left(\frac{z_1}{2}\right)\exp\left(-\frac{z_1\cos 2\theta}{2}\right)$$

$$= \exp\left[\frac{z_1}{2}\cos(2\theta+\pi)\right]\exp\left(\frac{z_1}{2}\right)$$

nd

$$\text{xp}\left[\frac{z_1}{2}\cos(2\theta+\pi)\right] = I_0\left(\frac{z_1}{2}\right) - 2I_1\left(\frac{z_1}{2}\right)\cos 2\theta + 2I_2\left(\frac{z_1}{2}\right)\cos 4\theta + \ldots \quad (44)$$

By substituting these relationships in eqn. (40) it is found that

$$R = a\varepsilon^{b/T}\left[I_0\left(\frac{b\theta_c}{T^2}\right) - 2I_1\left(\frac{b\theta_c}{T^2}\right)\sin\omega t - 2I_2\left(\frac{b\theta_c}{T^2}\right)\cos 2\omega t\right]$$
$$\times\varepsilon^{\left(\frac{b\theta_c^2}{2T^3}\right)}\left[I_0\left(\frac{b\theta_c^2}{2T^3}\right) - 2I_1\left(\frac{b\theta_c^2}{2T^3}\right)\cos 2\omega t + 2I_2\left(\frac{b\theta_c^2}{2T^3}\right)\cos 4\omega t\right] \quad (45)$$

Since only the mean value of the resistance is required, the eriodic terms can be omitted with the possible exception of the roduct term

$$+4I_1\left(\frac{b\theta_c^2}{2T^3}\right)I_2\left(\frac{b\theta_c}{T^2}\right)\cos^2 2\omega t \quad (46)$$

ecause this may contribute to the mean value of the resistance. The other terms, however, can be safely neglected, since the eries converges rapidly. The mean value of this product term is

$$2I_1\left(\frac{b\theta_c^2}{2T^3}\right)I_2\left(\frac{b\theta_c}{T^2}\right) \quad (47)$$

ince $\cos^2 2\omega t = \frac{1}{2}[1+\cos 4\omega t)$, and in consequence the mean esistance of the bead is

$$R = a\varepsilon^{\frac{b}{T}}\times\varepsilon^{\frac{b\theta_c^2}{2T^3}}\times\left[I_0\left(\frac{b\theta_c}{T^2}\right)I_0\left(\frac{b\theta_c^2}{2T^3}\right) + 2I_1\left(\frac{b\theta_c^2}{2T^3}\right)I_2\left(\frac{b\theta_c}{T^2}\right)\right] \quad (48)$$

ince the mean d.c. resistance at a steady temperature is $R = a\varepsilon^{b/T}$, it follows that the rest of the expression is a measure f the integration error.

The contribution due to the term

$$2I_1\left(\frac{b\theta_c^2}{2T^3}\right)I_2\left(\frac{b\theta_c}{T^2}\right)$$

is less than 20 parts in 10^6 for temperature variations less than $\pm 10°$ C, and is therefore entirely negligible. Hence, at low frequencies, the mean d.c. resistance at a steady temperature T changes by a factor A given by

$$A = \varepsilon^{\left(\frac{b\theta_c^2}{2T^3}\right)}\times\left[I_0\left(\frac{b\theta_c}{T^2}\right)\cdot I_0\left(\frac{b\theta_c^2}{2T^3}\right)\right] \quad (49)$$

(12.2) Errors Due to the Increase in Radiation Loss, when Temperature Fluctuations are Present

The heat loss by radiation from a surface is given by $\eta E(T^4 - T_0^4)$, where T is the surface temperature and T_0 the ambient temperature in degrees absolute. η is the Stefan–Boltzmann constant, and E is the emissivity of the surface. In general, for an oxide surface, the value of E lies between 0·6 and 0·7. If we insert numerical values it will be found that

$$\text{Loss of heat} = 5\cdot 75\times 10^{-12}\times E(T^4 - T_0^4)\ \text{watts/cm}^2 \quad (50)$$

If T is constant there will not be any frequency effect, but at low frequencies fluctuations will appear in T and this will cause a reduction in the mean temperature owing to increased radiation losses. These increased losses can be assessed in the following manner.

Assume that the instantaneous temperature of the whole bead surface is given by

$$T = \theta_c\sin\omega t + T_M \quad (51)$$

where T_M is the mean temperature.

The rate of loss of heat is

$$\eta E[(\theta_c\sin\omega t + T_M)^4 - T_0^4] \quad (52)$$

Expansion of $(\theta_c\sin\omega t + T_M)^4$ gives

$$\theta_c^4\sin^4\omega t + 4\theta_c^3 T_M\sin^3\omega t + 6\theta_c^2 T_M^2\sin^2\omega t + 4\theta_c T_M^3\sin\omega t + T_M^4 \quad (53)$$

The odd powers of ωt have an average value of zero over any number of complete cycles. The average value of $6\theta_c^2 T_M^2\sin^2\omega t$ is $3\theta_c^2 T_M^2$ and that of $\theta_c^4\sin^4\omega t$ is $3/8\theta_c^4$. Thus, the heat lost in any time t, where t contains an integral number of cycles, is

$$\eta E(T_M^4 + 3T_M^2\theta_c^2 + \tfrac{3}{8}\theta_c^4 - T_0^4)t \quad (54)$$

and the increase in heat loss due to temperature fluctuations is

$$3\eta E T_M^2\theta_c^2\left(1+\frac{\theta_c^2}{8T_M^2}\right) \quad (55)$$

The term $\theta_c^2/8T_M^2$ is small and can be neglected. If we take E as 0·7 and insert numerical values, we find that the increase in power loss is given by

$$\delta P = 12\cdot 1\times 10^{-12}a_2 T_M^2\theta_c^2\ \text{watts}$$

[The discussion on the above paper will be found on page 707.]

621.317.733 : 621.335.3

Paper No. 1963 M
Dec. 1955

A BRIDGE FOR THE MEASUREMENT OF PERMITTIVITY

By A. M. THOMPSON, B.Sc.

(*The paper was first received* 22*nd July, and in revised form* 17*th October*, 1955. *It was published in December*, 1955, *and was read before the* MEASUREMENT AND CONTROL SECTION 17*th April*, 1956.)

SUMMARY

A description is given of a null method for the measurement of the complex permittivity of dielectrics. The direct admittance of a 3-terminal capacitor with the sample as dielectric is measured as a complex capacitance, the two components being indicated directly on two 3-terminal variable air capacitors. In addition to these the bridge network comprises transformer ratio-arms and an amplifier whose output voltage is in quadrature with that of the transformer. The bridge operates at ten fixed frequencies from 30 to 10^6 rad/sec.

LIST OF SYMBOLS

C = Capacitance.
C_s = Complex capacitance of sample = $C' - jC''$.
L = Inductance.
R = Resistance.
Y = Admittance.
ϵ = Complex relative permittivity = $\epsilon' - j\epsilon''$.
V = Alternating voltage.
ω = Angular frequency.
m = Attenuator multiplying factor.
A = Amplifier gain.
$\alpha, \beta, \gamma, \delta$ = Small numbers, $\ll 1$.

(1) INTRODUCTION

The alternating-current properties of dielectrics are most simply expressed in terms of a complex permittivity, and one of the problems associated with fundamental investigations of dielectrics is the determination of this permittivity over the widest possible frequency range. At low frequencies this is done by constructing a capacitor of known dimensions with the sample as dielectric, and measuring its admittance. The capacitance is usually small (of the order of 10–100 $\mu\mu$F), and for accurate measurements the sample capacitor is preferably constructed as a 3-terminal capacitor with guard electrodes and separately shielded leads.

Many a.c. bridge circuits can be adapted for the measurement of 3-terminal capacitors by balancing the admittances to earth,[1] but the double balance can be avoided if the unwanted admittances are shunted across bridge arms of sufficiently low impedance. In the bridge to be described, transformer ratio-arms of very low effective impedance are used. To obtain a convenient conductance balance an amplifier controlled by a feedback network supplies a voltage in quadrature with that of the transformer, and the two components of the sample admittance are indicated on two 3-terminal air capacitors. Both values are obtained as equivalent capacitances, and multiplication by an appropriate geometrical factor is all that is required to obtain the two components of the permittivity. The bridge operates at ten fixed frequencies from 30 to 10^6 rad/sec.

(2) PRINCIPLE OF THE METHOD

If a pair of conductors having capacitance C_0 in a vacuum is immersed in a dielectric of relative permittivity ϵ, the admittance of the capacitor so formed may be conveniently represented as that of a complex capacitance whose value is given by

$$C_s = \epsilon C_0$$

The admittance will be $\quad Y_s = j\omega C_s$

Putting $\quad \epsilon = \epsilon' - j\epsilon''$ and $C_s = C' - jC''$

we have $\quad C' = \epsilon' C_0, \quad C'' = \epsilon'' C_0$ and $Y_s = j\omega C' + \omega C''$

A simple way to measure such an admittance is to compare it with a known variable admittance in an equal-ratio bridge. The variable admittance should provide for independent adjustment of capacitance and conductance. A pure variable capacitance can be provided quite simply by a 3-terminal variable air capacitor, but the provision of a pure variable conductance of very small value for a wide frequency range is much more difficult. The variable conductance may be replaced by a variable capacitance if a voltage source in quadrature with that of the bridge supply is added to the bridge.

The ideal voltage distribution in the network so obtained is shown in Fig. 1. The terminal voltages are expressed with respect to a common reference which is the earthed point of th[e] system and to which the guards and shields are connecte[d]. The admittance to earth from the detector terminal is repre[-]sented by an equivalent capacitance C_3.

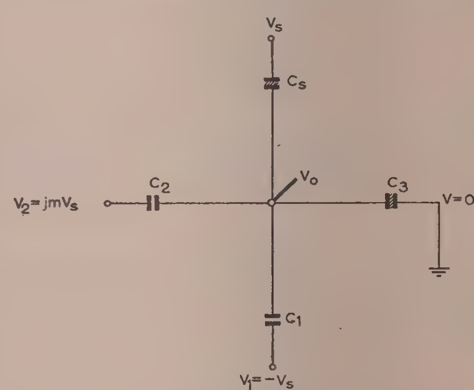


Fig. 1.—Network with ideal voltage distribution.

It follows from Millman's theorem[2] that

$$V_0 = \frac{V_s C_s + V_1 C_1 + V_2 C_2}{C_s + C_1 + C_2 + C_3}$$

If $V_1 = -V_s$ and $V_2 = jmV_s$, for $V_0 = 0$ we have

$$C_s = C_1 - jmC_2 \quad . \quad . \quad . \quad . \quad . \quad ($$

i.e. $C' = C_1$ and $C'' = mC_2$ and therefore

$$\epsilon' = \frac{C_1}{C_0} \quad \text{and} \quad \epsilon'' = \frac{mC_2}{C_0}$$

Mr. Thompson is at the Division of Electrotechnology, Commonwealth Scientific and Industrial Research Organization, Australia.

(3) MEASURING SYSTEM

The major components of a measuring system based on Fig. 1 are shown in Fig. 2. Voltages of opposite phase are obtained from a pair of secondary windings on a transformer and a voltage

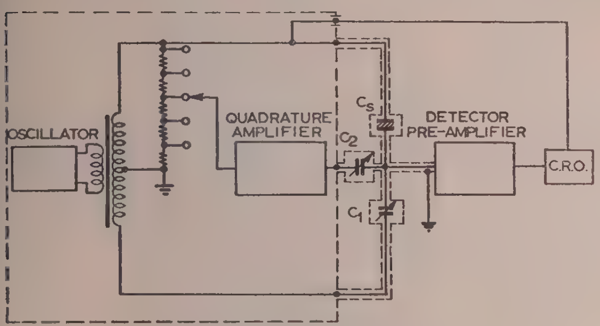


Fig. 2.—Components of measuring system.

in quadrature with these is obtained from a feedback amplifier whose input is derived from a stepped attenuator connected across one of the secondary windings. The sample and standard capacitors are connected to these voltage sources, and the junction voltage V_0 is amplified and indicated on a cathode-ray oscillograph. This voltage may be reduced to zero by adjusting the capacitors C_1 and C_2 and the attenuator, which has five decade steps ($m = 1 \rightarrow 10^{-4}$). The components of the sample capacitance are read directly from the scales of C_1 and C_2, the latter being multiplied by the setting of the attenuator.

(4) ACCURACY REQUIREMENTS

To examine the effect of imperfect voltage ratios let

$$V_1 = -(1 + \alpha + j\beta)V_s \quad \text{and} \quad V_2 = jm(1 + \gamma + j\delta)V_s$$

Then eqn. (1) becomes

$$C_s = (1 + \alpha + j\beta)C_1 - jm(1 + \gamma + j\delta)C_2$$

$$C' = C_1\left(1 + \alpha + \delta\frac{mC_2}{C_1}\right), \text{ and } C'' = mC_2(1 + \gamma) - \beta C_1$$

It was desired that the bridge should be usable without corrections to an accuracy of 1% for C' and C'', with an absolute limit on C'' of $10^{-4}C'$. To meet these requirements $|\alpha|$ must be $<10^{-2}$ and $|\beta| <10^{-4}$. The limits for γ and δ depend on the tangent of the loss angle of the dielectric, mC_2/C_1, which may be between 1 and 10^{-4}, but they are in all cases satisfied for $|\gamma|$ and $|\delta| < 10^{-2}$. The most critical requirement is for β, the phase angle between the two transformer voltages.

(5) SOURCES OF ERROR

(5.1) Transformer Ratio-Arms

The ratio of the secondary voltages of the transformer may differ from the turns ratio owing to two separate effects. The voltages induced in the secondary windings may be unequal because of leakage flux which does not link both, or the loads on the windings may be sufficiently unbalanced to produce an error. The loading error is due to the resistance and leakage inductance of the loaded secondary winding, and may be determined by setting up the bridge and measuring the change in ratio for known additional loads on each winding. When the loading effects are known, the open-circuit ratio may be determined by reversal of the ratio windings with respect to the rest of the bridge.

The equivalent circuit of the ratio transformer is shown in Fig. 3, where the transformer is represented by an ideal trans-

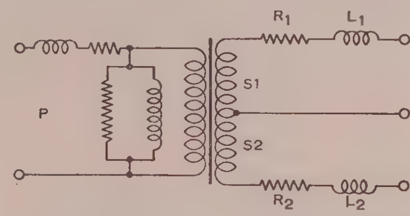


Fig. 3.—Equivalent circuit of ratio transformers.

Low frequency $R_1 \simeq R_2 = 160$ ohms; $L_1 \simeq L_2 = 3{\cdot}8$ mH.
High frequency $R_1 \simeq R_2 = 0{\cdot}4$ ohm; $L_1 \simeq L_2 = 1{\cdot}8\,\mu$H.

former and leakage impedances. Only the ratio of the two secondary windings and the magnitudes of the secondary leakage impedances which determine the loading errors need be considered. Two transformers were used to cover the frequency range, and in both of them the error in the open-circuit ratio was less than 2×10^{-5}. An unbalanced load of 200 $\mu\mu$F would cause an error of phase angle of 10^{-4}radn at the highest frequencies in both transformers but there is usually no difficulty in ensuring that the unbalanced load remains smaller than this. The load due to the attenuator (about 10 kΩ) is balanced by a similar resistive load on the other half of the transformer.

(5.2) Quadrature Amplifier

The quadrature amplifier consists of a voltage amplifier with a flat frequency-response and an RC feedback network as shown

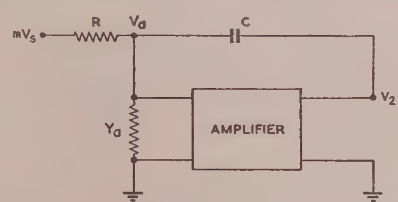


Fig. 4.—Schematic of quadrature amplifier.

in Fig. 4. Writing A for the gain of the amplifier and Y_a for its input admittance, we have

$$V_a = \frac{V_2}{A} = \frac{mV_s/R + V_2 j\omega C}{1/R + j\omega C + Y_a}$$

whence

$$V_2 = \frac{jmV_s}{\omega CR + j(1 + j\omega CR + Y_a R)/A}$$

The components C and R are switched for frequency changes, so that $\omega CR = 1$. Under these conditions V_2 is approximately equal to jmV_s, and since $|Y_a| \ll \omega C$ the error in this approximation is of the order of $1/A$ in both magnitude and phase, and will be less than 1% for $|A| > 100$. Any errors in the attenuator or in the components C and R will add to this error. However, the attenuator is not used when $m = 1$ and it was found that with careful layout no additional phase compensation was necessary to keep the contribution $\delta(mC_2/C_1)$ under 1%.

(6) CONSTRUCTION

The oscillator, transformer ratio-arms and quadrature amplifier were built as a single unit, so that ganged controls could be

used to switch transformers and components for frequency changes. The sample holder and the two 3-terminal air capacitors are connected to this supply unit and to a separate detector pre-amplifier by coaxial cable. Simplified circuit diagrams of the oscillator, quadrature amplifier and detector pre-amplifier are shown in Figs. 5, 6 and 7, respectively.

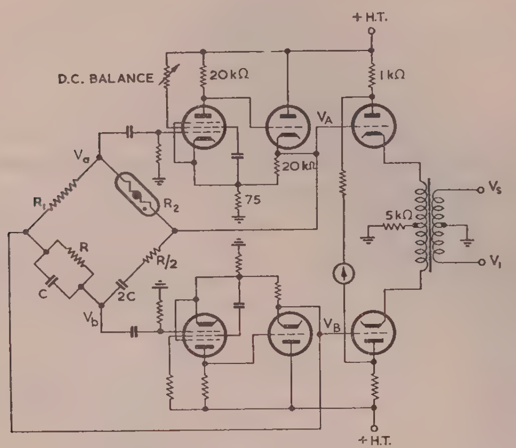


Fig. 5.—Oscillator circuit.

Nominal operating conditions

$$\omega = 1/RC; \quad R_1 = R_2$$
$$V_A = V_B \quad V_a, V_b \to 0$$

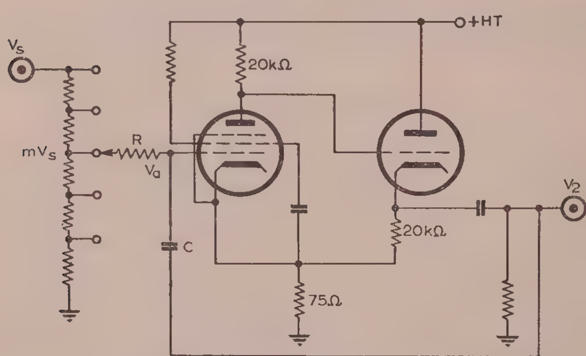


Fig. 6.—Quadrature-amplifier circuit.

Nominal operating conditions

$$\omega CR = 1; \quad V_2 = jmV_s; \quad V_a \to 0$$

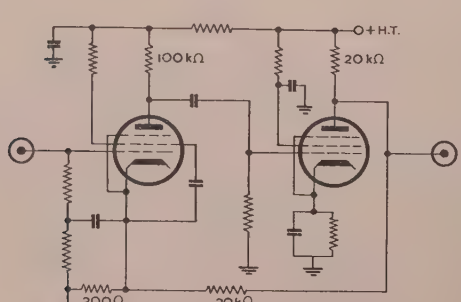


Fig. 7.—Detector pre-amplifier circuit.

Gain = 100.

(6.1) Oscillator

The bridge-stabilized oscillator incorporates a Wien bridge for frequency control and a directly heated thermistor for amplitude stabilization. Two separate amplifiers are used in the maintaining circuit to provide a broad-band drive to the Wien bridge, which is not directly earthed. The frequency is switched in fixed steps by varying the capacitors in the bridge, and a 2-position range switch is used to switch resistors and at the same time to select the appropriate output transformer. The oscillator is directly coupled to a pair of cathode-followers which drive the output transformer. It is necessary to balance the direct currents in the output valves to avoid d.c. magnetization of the transformer and hence distortion at low frequencies. The out-of-balance current is indicated on a small meter. The adjustment to balance also reduces the direct current in the thermistor to a level at which it does not affect the amplitude of oscillation.

(6.2) Quadrature Amplifier

The three amplifiers, two for the oscillator and one for the quadrature voltage, are identical and consist of a broad-band pentode stage followed by a direct-coupled cathode-follower. Cathode bias for the pentode is obtained from a resistor which is common to both cathode circuits. Since the cathode load of the output valve is the same as the anode load of the pentode, alternating currents cancel in the bias resistor and in the power supply. The gain of the amplifier in the operating frequency-range is about 140. For the quadrature amplifier an additional capacitance coupling is necessary to obtain an output with zero direct voltage with respect to earth.

(6.3) Transformers

The low-frequency transformer is wound on a silicon-iron core and the secondary windings consist of two stacks each of four coils. Alternate coils in each stack are connected in series. By this means a symmetrical transformer with low self-capacitance is obtained. The high-frequency transformer is wound on a Radiometal core, a twisted-bifilar winding constituting the two secondary windings.

(6.4) Three-Terminal Capacitors

The measuring capacitors are external to the supply unit and different capacitors have been used from time to time. For general use two 3-terminal capacitors were constructed from variable air capacitors by mounting both sets of plates separately on a metal base. The insulators used—ceramic washers—are inset so that they do not contribute to the direct capacitance, which is then practically loss-free.[3] The scales are approximately logarithmic and can be read within 1% over the range 10–100 $\mu\mu$F. Slow-motion drives are fitted and the bridge may be balanced to 1 part in 10^4 as required for the measurement of low power factors.

(6.5) Detector

The sensitivity required of the bridge is of the order of 0·001 $\mu\mu$F, which corresponds to a few tens of microvolts at the detector. A pre-amplifier with a flat frequency-response and a gain of about 100 has been used successfully, the output being displayed on a small standard oscillograph with a sinusoidal time-base supplied by the bridge oscillator. Provided that the input conductance of the pre-amplifier is small enough, the overall sensitivity is independent of frequency. The pre-amplifier consists of two pentode stages with feedback to give a flat frequency-response and a small input conductance.

(6.6) Power Supply

A common power supply is provided as a separate unit. The h.t. supply is regulated and a rectified filament supply is included. This filament supply is used for the pre-amplifier and the first valves of the oscillator and quadrature amplifiers, and results in a considerable reduction in 50 c/s interference compared with an a.c. supply.

(7) PERFORMANCE

The oscillator frequency and the output of the quadrature amplifier depend on the stability of a number of components, namely high-stability carbon resistors and mica capacitors. It was found that the heat dissipated in the instrument caused a drift in some components of $\pm\frac{1}{2}\%$, and limited the accuracy of adjustments to a range of this order. However, this accuracy is sufficient to meet the requirements laid down in Section 5. The overall performance of the bridge was checked by measuring known combinations of capacitors and resistors in parallel and was found to be within the specified limits. T-networks were used to obtain the equivalent of the very small conductances required for tests at the lower frequencies.

(8) CONCLUSION

The bridge has been used for a number of years for measurements on a wide range of dielectrics, and has proved very convenient and rapid in operation. The oscillator and quadrature-amplifier frequency controls being ganged, operation of the bridge is independent of frequency. This enables the sample to be measured as a complex capacitance and reduces computation to a minimum. A major advantage of the bridge is that it enables a 3-terminal capacitor to be measured without the need for a separate balance of the earth admittances.

(9) REFERENCES

(1) Harris, F. K.: "Electrical Measurements" (Chapman and Hall, 1952), p. 730.
(2) Millman J.: "A Useful Network Theorem," *Proceedings of the Institute of Radio Engineers*, 1940, **28**, p. 413.
(3) Astin, A. V.: "Nature of Energy Losses in Air Capacitors at Low Frequencies," *National Bureau of Standards Journal of Research*, 1939, **22**, p. 673.

DISCUSSION ON THE ABOVE TWO PAPERS BEFORE THE MEASUREMENT AND CONTROL SECTION, 17TH APRIL, 1956

Dr. A. H. M. Arnold: With reference to the paper by Mr. Widdis I would like to know whether the accuracy of $2\frac{1}{2}$ parts in 10 000 rests purely on the final test of the thermistor against the thermocouple. It may be that the thermocouple was not as good as the author thought, since I would have credited the thermistor with an accuracy of 1 part in 10 000.

The thermistor may prove very valuable as an a.c. precision transfer device, but its long time-constant would make it very inconvenient for standardizing an a.c. potentiometer.

I would like first to consider the measurement of alternating current without the use of thermistors. Primarily, of course, we rely on the dynamometer and have done so for 50 years. It is a very good instrument—or it can be. I know it can also be a bad one, but it has two advantages over the thermistor—it just reads alternating current or direct current without any intercomparison between the two, and it has a quick response. Some 30 years ago a dynamometer was made, for use with potentiometers by a firm with which the author has had some connection, which had an a.c.–d.c. transfer error of something less than 1 part in 10 000. I regret that it is no longer made. However, the dynamometer has its limitations in the upper and lower frequency ranges. The electrostatic voltmeter can cover a higher frequency—up to 100 kc/s or possibly higher.

The middle frequency range is very adequately covered by the dynamometer and the electrostatic voltmeter, and thus there are three possible uses for the thermistor. The first is to check the dynamometer and electrostatic voltmeter. Neither of these instruments is perfect, and we cannot have too many checks. The second is to provide a transfer device in the frequency ranges where those instruments cannot be used. In the lower frequency range we have to admit that the thermistor has no competitor. In spite of what the author states, I think that the long time-constant is very inconvenient, but if measurements have to be made at 1 c/s or lower, there is no help for it, and the thermistor will do the job better than anything else.

On the other hand, at the higher frequencies, I am not so convinced that the thermistor can hold its own against the thermocouple. It is true that the d.c. errors on the thermocouple due to Peltier and Thomson effects may be troublesome. There is not much error with nickel–chromium heaters when the mean of the two polarities is taken, but a large reversal error is a nuisance, and there may be a small residual error. This can be checked quite well, of course, at low frequencies; and in that case the thermistor might be useful for calibrating the thermocouple. But at the high-frequency end, the a.c. errors of the thermistor would tend to be higher than those of the thermocouple. The heater is in the form of a coil, which increases the capacitance and also the eddy currents, so that the thermocouple with its short straight heater wire should be better.

So far as sensitivity is concerned, I do not attach a great deal of value to the author's claim of some large factor by which the thermistor gains, because this advantage would be negatived by the long time-constant. At any rate, there is no difficulty in obtaining an accuracy of about 1 part in 10 000 with a thermocouple, which is quite adequate for most purposes at present.

Dr. C. J. N. Candy: With reference to the paper by Mr. Widdis, I have used indirectly heated thermistors as transfer elements in very-low-frequency circuits. The non-linearities and the thermal drift were partly eliminated by using two thermistors in a push-pull circuit. The beads were connected in series and polarized in opposite directions, so that the standing voltage across the combination was zero; the heaters were connected in a similar way. The signal was applied to both heaters in series, so that the net current in one increased as the other decreased.

In Section 3 a method for determining the thermistor parameters is presented. The process of heating the thermistor in an oil bath may be avoided by making measurements on the voltage/current characteristic alone. It follows from eqn. (1) that

$$\frac{1}{\log(R/R_0)} = -\frac{T_0^2 K}{b}\left(\frac{1}{P}\right) - \frac{T_0}{b}$$

where R_0 is the resistance of the thermistor at the ambient temperature T_0. A straight line is obtained by plotting $1/\log(R/R_0)$ against $1/P$; the intercepts with the axis may be used to calculate b and K; a is determined by substituting R_0 and T_0 in eqn. (1).

Mr. A. Felton: The thermistor is a comparatively recent development, and it owes its rapid advance as a precision instrument in no small degree to the simultaneous development of electronically stabilized supplies. Ten years ago it would have been of little

value as a transfer device, because its long thermal time-constant would have made it excessively difficult to use on supplies from rotating machines or from the imperfectly stabilized oscillators of those days.

This long time-constant is a unique feature of the thermistor, and while it is turned to advantage in the measurement of very low frequencies, it is a handicap in the great majority of applications. Nevertheless, I feel sure that the thermistor will be used more and more in measurement.

My only criticism of the paper by Mr. Widdis is that, in his enthusiasm for the new device, he has done something less than justice to the old.

Thus I cannot agree with his description of the electrostatic voltmeter as elaborate and unsuitable for use outside a national standardizing laboratory. The voltmeter as used at the National Physical Laboratory has remained unchanged since Rayner developed it from the Kelvin instrument before 1920. It is perfectly satisfactory without modification up to 100 kc/s, and it is inherently simple to make and install. The auxiliaries, too, are no more elaborate than those used with any other precision transfer instrument; and several commercial firms, as well as Commonwealth and other standardizing laboratories, have adopted it with every satisfaction.

With regard to the dynamometer, it is true that the power consumption is rather high, but this is not always a disadvantage. For example, a transfer instrument is very often used solely to calibrate an a.c. instrument, and in that case the power consumption—so long as it is within the capacity of the generator to supply it—is quite immaterial.

The thermal convertor is a well tried transfer device and one likely to find increasing utility. It is true that it is temperature dependent, but so is the thermistor, and precautions used with the one can be applied to the other. It is true also that a certain amount of selection is required, but are all samples of thermistors equally suitable for use as transfer standards?

With regard to the performance of thermistors at high frequencies, the author calculates that the error should be 1 part in 10^4 at 500 kc/s. We have found that errors of this order occur at only 100 kc/s, which is a discrepancy between theory and practice that might be worth following up.

Dr. L. Hartshorn: With reference to the paper by Mr. Thompson, as a research tool for the insulation laboratory the present bridge has much to recommend it. What appeals to me most is that the quantities observed are two capacitances, one determining the permittivity, and the second one divided by the first one determining $\tan\delta$. This is essentially the same condition that W. H. Ward and I obtained in our resonance method for permittivity and $\tan\delta$ in the region 10^4–10^8 c/s, and I have always believed that the widespread use of our method was in large measure due to that feature, because it means that the only calibrations required are those of two simple scales which serve for all frequencies.

It is interesting that the paper is in the direct line of two earlier contributions to the *Proceedings*, which were themselves based on the still earlier work of Blumlein on the one hand and C. G. Mayo on the other. The theme has been variously described as the transformer bridge, developed in the radio range largely by C. G. Mayo and described by Kirke* in 1945 in the Chairman's Address to the Radio Section, and then as inductively-coupled ratio-arms by Clark and Vanderlyn,† who described developments in audio-frequency bridge techniques that began with Blumlein's patent specification of 1928. The advantages of a Wagner earth are secured without the complication of a double balance. Clark and Vanderlyn used the scheme to measure direct impedances in networks in which the indirect impedances would, if uncorrected, amount to short-circuits, while Mayo used it for bridge measurements of ordinary impedances at frequencies up to 100 Mc/s.

These techniques seem to have a future, and the bridge is not only an ingenious and useful research tool but another good example of the importance of Blumlein ratio arms.

Mr. J. K. Webb: Dr. C. J. N. Candy* recently presented two papers on the thermistor, and the paper by Mr. Widdis now illuminates another facet of its usefulness. In the same context, mention might be made of a very serviceable instrument using thermistors for measuring the r.m.s. values of small currents from zero to video frequencies.†

The short paper by Mr. Thompson is a model of clarity, although the author might have avoided quoting angular frequencies, and mentioned that $\tan\delta = mC_2/C_1$, where δ is the dielectric loss angle. His bridge has the merit of incorporating components with which we have now become familiar, such as transformer ratio arms and feedback amplifiers, in an original manner. On the face of it, the arrangement appears to be eminently practical, but experience has nevertheless taught me that, however simple and straightforward a circuit appears on paper, unexpected difficulties may arise before a working model becomes commercially available. A Schering bridge which, despite its drawbacks, will do all that Mr. Thompson's bridge purports to do, and perhaps more, has been developed to this stage. It was based on an N.P.L. design of Dr. T. I. Jones and incorporates coaxial-type resistors, designed by Dr. N. F. Astbury, as ratio arms. Difficulties with the Wagner balance can be minimized by means of the cathode-follower device described by Rayner and Willner.‡ This bridge has given, and is giving, such good service that before we turn to any alternative scheme we should be quite certain that it is really worth while. I hope, therefore, that the author can get his bridge into production so that we shall then better be able to assess its virtues.

Mr. A. C. Lynch (*communicated*): The use of inductively-coupled ratio arms is slightly shocking to anyone accustomed to the older types of components for bridge networks, but they have real advantages which ought not to be overlooked. Their main drawback is that reactance must be balanced by reactance, and resistance by resistance; the balancing of a loss by adjusting a capacitance, so convenient a feature of the Schering bridge, is impossible. The bridge described in the paper by Mr. Thompson overcomes this drawback in a new way, the effectiveness of which can be judged only by experience. As it stands, it can be used only at a set of fixed frequencies, but I think there might be a way of deriving a quadrature voltage, even for a continuously-variable frequency, by using a heterodyne oscillator with phase-shifting at the primary frequency and two independent mixers.

By having the specimen and the variable capacitor in opposite arms of the network, rather than using a substitution method in which they would be in the same arm, the author has made the design of his transformers more difficult. I have a paper in preparation describing a substitution bridge in which any ordinary transformers can be successfully used.

* Kirke, H. L.: Radio Section: Chairman's Address, *Journal I.E.E.*, 1945, 92, Part I, p. 39.
† Clark, H. A. M., and Vanderlyn, P. B.: 'Double-Ratio A.C. Bridges with Inductively-Coupled Ratio Arms', *Proceedings I.E.E.*, Paper No. 742 M, January, 1949 (96, Part II, p. 365).

* Candy, C. J. N.: 'The Specification of the Properties of the Thermistor as a Circuit Element in Very-Low-Frequency Systems', *Proceedings I.E.E.*, Paper No. 1751 M, December, 1954 (103 B, p. 398).
Candy, C. J. N.: 'A Vector Method for Amplitude-Modulated Signals', *ibid.*, Paper No. 1772 M, January, 1955 (103 B, p. 410).
† Wood, H. B.: 'An R.M.S. Milliammeter of Novel Design for the Measurement of Current from Zero to Video Frequencies', *Journal of Scientific Instruments*, 1954, 31, p. 124.
‡ Rayner, C. H., and Willner, R. W.: 'A Method of Decreasing the Effect of Earth Admittance in A.C. Bridges', *Journal of Scientific Instruments*, 1950, 27, p. 103.

THE AUTHORS' REPLIES TO THE ABOVE DISCUSSION

Mr. F. C. Widdis (*in reply*): I am in general agreement with Dr. Arnold's comments. The accuracy of the thermistor is probably one part in 10 000. My original estimate was based upon checks against vacuo-thermo-junctions, and I have recently carried out a reassessment of the errors which indicates that I was over pessimistic. The large time-constant of the thermistor is a disadvantage if very-low-frequency measurements are not envisaged. I feel, however, that, with the very minute thermistor beads now available, it may be possible to construct a transfer device using them with a time-constant comparable with that of the vacuo-thermo-junction. The main disadvantage of the vacuo-thermo-junction, apart from the question of selection, appears to be the thermal drift associated with it. I believe that this drift is largely due to the existence of unwanted thermo-junctions in the thermocouple leads, and the use of balanced pairs for drift compensation is therefore not very successful. It is possible that alterations in the construction of the actual thermocouple might eliminate this defect, but at present there appear to be some practical difficulties in effecting such alterations.

In reply to Mr. Felton, I must apologize if I have been unjust to other devices; this was not my intention. I feel that only a limited number of laboratories would require the installation of permanent equipment for standardizing purposes, and any simple device which could be used for this purpose deserves consideration. The thermistor has the advantage shared with the vacuo-thermo-junction that it is inexpensive and readily available and can possibly be used with existing equipment when the need arises. It must be looked upon simply as an additional tool.

All the thermistors I examined had low reversal differences and selection was not necessary. Difficulties might occur in achieving temperature compensation if by any chance a pair of thermistors with substantially differing values of b were used.

I am surprised at the discrepancy between theory and practice at high frequencies; this deserves further investigation.

The suggestion put forward by Dr. Candy would considerably simplify the determination of the thermistor parameters.

Mr. J. K. Webb drew my attention to the r.m.s. milliammeter designed by H. B. Wood some years ago, and this largely prompted my investigation into the performance of thermistors.

Mr. A. M. Thompson (*in reply*): As Dr. Hartshorn points out, one of the major advantages of transformer ratio arms is that they enable direct admittances to be measured accurately without the need of a Wagner earth and consequent double balance. This application was suggested by G. A. Campbell* in 1922 although the differential transformer was used in bridge circuits very much earlier than this.

In reply to Mr. Webb I should say that, for the purpose for which it was designed, this bridge is very much more convenient to use than a Schering bridge. The absence of a Wagner earth and the direct-reading feature reduce considerably the time required to determine the properties of a dielectric at a number of different frequencies.

As suggested by Mr. Lynch, it is certainly possible to derive a two-phase supply of sufficient amplitude and phase stability to obtain a measuring circuit with a continuous variation of frequency. The substitution bridge he has developed is very effective in eliminating errors in the transformers. However, for the frequency range considered, the design of the transformers presents no great difficulty and the double balance is not necessary.

* CAMPBELL, G. A.: 'Measurement of Direct Capacities', *Bell System Technical Journal*, 1922, 1, p. 18.

DISCUSSION ON

'THE APPLICATION OF THE HALL EFFECT IN A SEMI-CONDUCTOR TO THE MEASUREMENT OF POWER IN AN ELECTROMAGNETIC FIELD'*

AND

'THE DESIGN OF SEMI-CONDUCTOR WATTMETERS FOR POWER-FREQUENCY AND AUDIO-FREQUENCY APPLICATIONS'†

Before the North-Eastern Radio and Measurement Group *at* Newcastle upon Tyne 16*th January, and the* Western Centre *at* Bristol 12*th March*, 1956.

Dr. E. E. Schneider (*at Newcastle upon Tyne*): I am interested in the application of the method to the measurement of power at microwave frequencies and in the ingenious device, for use in waveguides, which the author describes. However, since it employs a resonant cavity, it will be very frequency sensitive. Is the author also developing more broad-banded devices, which could be used for continuously monitoring the power flow in waveguides?

Prof. J. C. Prescott (*at Newcastle upon Tyne*): The possibility of using a germanium crystal as a multiplying device might well be exploited in the design of a harmonic analyser for current or voltage waves of high frequency and low energy content. It is possible to design an analyser for low-level high-frequency work if we are prepared to use an electrometer, but this requires very careful adjustment and is not easily transportable. If, however, a germanium crystal were used, the wave to be analysed could provide the p.d. between two faces of the crystal while the transverse magnetic field was generated by an oscillator, the frequency of which could be varied over the spectrum of the harmonics to be sought. With a d.c. voltmeter connected to the crystal to detect the Hall effect, steady readings would appear whenever the oscillator frequency coincided with one of the harmonic components of the wave to be analysed; some method would probably be desirable to anchor the oscillator more or less rigidly to this wave, but I do not think this need present any serious difficulty.

Mr. W. F. Collinson (*at Bristol*): How is the calibration of Hall-effect wattmeters verified for use at frequencies of the order of hundreds of megacycles?

Mr. B. F. Hallett (*at Bristol*): The Hall-effect voltage is proportional to the product of the current, i_c, through the crystal and the flux density, B, of the magnetic field crossing it at any instant. If this voltage is to be proportional to the instantaneous power in the associated circuit, there must be no phase displacement between i_c and the electric field, and between B and the magnetic field of the circuit under test. In view of the wide range of frequencies at which this principle has been used, what difficulties have been met in preventing this trouble and how have they been overcome?

Mr. G. O. McLean (*at Bristol*): Although the supply industry welcomes the author's proposal to measure power at 132 kV without the use of current transformers, a greater economic benefit would be obtained from the elimination of the voltage transformers. What are the prospects of achieving this?

Mr. A. H. McQueen (*at Bristol*): The author describes an arrangement whereby the magnetic field surrounding a bushing might be used to operate a wattmeter to measure the power flow through the bushing. With the present arrangement using current and voltage transformers or capacitance bushings, certain accuracies are defined by British Standards. Would the same accuracies be obtained with the author's arrangement?

Prof. H. E. M. Barlow (*in reply*): The frequency bandwidth over which the resonant-cavity arrangement can be used for microwave power measurement depends primarily on the Q-factor of the cavity with its associated semi-conductor element; when this is low an ample frequency band is easily obtainable, but the power absorbed by the instrument as a wattmeter is then correspondingly higher. A compromise is clearly necessary, but I can assure Dr. Schneider that there is no great difficulty in obtaining a reasonable performance. The transmission type of instrument employing a semi-conductor element mounted directly in the main guide is capable of operation over a wider frequency band, but the improvement in this respect is obtained at the expense of sensitivity to lower powers.

Professor Prescott's suggestion offers another interesting application of the 'Hall effect' instrument, and provided that rectifier action is substantially eliminated, I see no reason to doubt that a practical harmonic analyser could be constructed in the way he outlines.

As Mr. Collinson will appreciate, the semi-conductor wattmeter described is not an absolute one, but its calibration for u.h.f. use can be conveniently carried out by converting the power to be measured, or a known proportion of it, into heat, thus enabling a direct comparison to be made with a corresponding conversion at low frequency.

Mr. Hallett has drawn attention to the importance of maintaining the correct phase relationship between the electric and magnetic fields acting on the semi-conductor element. At low frequencies this is not difficult to achieve, because it is sufficient to establish the same ratio of resistance to inductance for the two circuits of the wattmeter, as described in the second paper. In the microwave application, however, the semi-conductor element carries a displacement component of current having the same order of magnitude as the conduction component on which the true Hall e.m.f. depends. The behaviour in this case is still under investigation, but without knowing precisely what happens within the semi-conductor it has, in fact, proved possible to make a satisfactory working instrument.

The elimination of the voltage transformer as an essential part of power-measuring equipment, which Mr. McLean points out would be particularly helpful, might be achieved either by using an h.v. insulator or bushing as a capacitance voltage-divider or by relying upon a high-resistance leak circuit. Only a few milliamperes of current are required to feed the semi-conductor element, and the problem therefore does not seem to be insuperable. Experiments are in progress to ascertain the possibilities in applying the semi-conductor device to an h.v. bushing, but the work is not yet sufficiently advanced to give Mr. McQueen a reliable estimate of the accuracy obtainable from this kind of wattmeter.

* Barlow, H. E. M.: Paper No. 1654 M, June, 1954 (see 102 B, p. 179).
† Barlow, H. E. M.: Paper No. 1778 M, November, 1954 (see 102 B, p. 186).

621.396.99:523.164

Paper No. 2227
Nov. 1956

THE FORTY-SEVENTH KELVIN LECTURE

'RADIO ASTRONOMY AND THE JODRELL BANK TELESCOPE'

By Prof. A. C. B. LOVELL, O.B.E., M.Sc., Ph.D., F.R.S.

(*Lecture delivered before* THE INSTITUTION, 26*th April*, 1956.)

SUMMARY

The lecture describes some of the scientific characteristics of the 250 ft-aperture steerable radio telescope at the Jodrell Bank Experimental Station of the University of Manchester. Some of the research problems on which the telescope will be used are then outlined against the current background in radio astronomy. These include the study of the background continuum radio emission of the Galaxy and of the galactic radio sources over a wide range of wavelengths. The main programme on the extragalactic radio emissions will involve a study of the spatial distribution of particular classes of extragalactic radio sources and the measurement of their distances, from which it is hoped that significant progress can be made with cosmological problems. The telescope will also be used as a combined transmitter and receiver to make further studies of the moon and possibly the planets. Other items in the programme include the study of very faint meteors by the radio echo technique and various problems of solar terrestrial relationships of particular importance to the International Geophysical Year.

(1) INTRODUCTION

There are grounds for hoping that the year of the 47th Kelvin Lecture may see the beginning of the initial tests and experiments with the fully steerable 250 ft-aperture radio telescope at Jodrell Bank. It may therefore be appropriate to describe some of the scientific aspects of this instrument and the programme of work in radio astronomy on which it will be used.

The discovery by K. G. Jansky in 1932 that radio waves of extra-terrestrial origin were reaching the earth passed almost unnoticed. Even in 1950 when Jansky died we read that his brother,[1] himself an electrical engineer, expressed the opinion that the value of the discovery would not be recognized for at least half a century. Even so, at that time the need for telescopes of large aperture to pursue the investigation of these radio waves had been fully realized at Jodrell Bank. A large transit telescope of 218 ft aperture was already in operation, and with it Hanbury Brown and Hazard[2] had discovered that the extragalactic nebulae were sources of radio emission as well as the milky way system. Both in this aspect of radio astronomy and in the radar or radio-echo work, a fully steerable telescope of similar aperture seemed essential for the further exploration of space. Indeed by 1950 the scientific design of such a telescope had developed to the stage when the support of the Royal Astronomical Society could be enlisted. There remained the difficult problems of finance and engineering. The Department of Scientific and Industrial Research had, at an early stage, indicated its interest in the proposal, and in the spring of 1952 the Department announced that, with the assistance of the Nuffield Foundation, it was able to bear the cost of construction. Messrs. Husband and Co., of Sheffield and London, were appointed the engineers for the project, and the work on the site was initiated in the following autumn.

It is not possible in this lecture to refer to the many difficult and unprecedented problems which faced the engineers, and my intention is limited to a description of the anticipated scientific performance of the telescope and of the immediate problems on which it will be used.

(2) SOME SCIENTIFIC ASPECTS OF THE RADIO TELESCOPE

(2.1) General Considerations

At the time when the fundamental decisions had to be made about the type of radio telescope to be built the issues were considerably simpler than they are to-day. In particular, the existence of the 21 cm interstellar hydrogen line was little more than a theoretical speculation, and in fact no galactic or extragalactic radio emissions had been detected at wavelengths much below 1 m. Similarly, in the radio-echo work almost all research had been carried out on wavelengths above 1 m. Hence there was no justification for gambling a very large sum of money on an instrument in an untried wavelength range, when so many problems of absorbing interest and importance awaited solution in the waveband from about 1 to 15 m. Further, the existing experience with the 218 ft transit telescope indicated that a steerable telescope with at least an equivalent aperture was needed on grounds of power gain and resolution, to pursue the pressing problems which were opening up in all aspects of radio astronomy.

The final choice was for a full paraboloid of aperture 250 ft with the focus in the plane of the aperture. The paraboloid was an obvious choice of aerial system for an instrument in which maximum ease of change of wavelength is essential, and which has to be used as either a receiver or a transmitter. The choice of focal length presented more difficulty. In the 218 ft transit telescope the focal length is long, the primary feed being carried on a tower 126 ft above the ground. At the time, this was largely determined by the easier mechanical construction involved in a shallow bowl, but there is a severe loss of power gain because of the overspill radiation from the primary feed. On the other hand, in this transit telescope it has been easy to design a primary feed giving a uniform distribution of intensity over the aperture plane. In general, as the focal length of such a telescope is decreased the overspill is reduced but the uniformity of the polar diagram of the primary feed over the aperture plane decreases and the beam shape worsens. It can be shown that it is nearly always possible to design a primary feed so that, as far as power gain is concerned, these two effects nearly balance for changes in focal length, and hence the choice can be governed by other considerations. In the case of the steerable telescope the dominant consideration was that the power received in the overspill radiation should be reduced to the minimum possible. A further practical factor is that, to achieve optimum performance, the size and complication of the array used for the primary feed must increase with the focal length. Thus, increases in focal length demand a lengthening of the central tower carrying greater loads at the focus, which must be maintained accurately in all positions of the telescope—a consideration which again suggests the desirability of an instrument of short focal length. These kinds of criteria determined the ultimate choice of a focal plane design for the steerable telescope.

Prof. Lovell is Professor of Radio Astronomy in the University of Manchester and Director of the Jodrell Bank Experimental Station, Cheshire.

As originally designed, the reflecting material of the paraboloidal bowl was a mesh of stiff copper wire. Detailed investigations by the engineers, however, showed that it would be extremely difficult and costly to make the vast number of joints between the individual sheets of mesh, and the instrument is now to be surfaced with a solid membrane of welded steel sheets. This will be, of course, a considerable advantage when the telescope is operated on short wavelengths, since all losses due to transmission through the mesh will be avoided. Measurements of the electrical resistance on a sample section of the welded reflector have shown that for practical purposes the reflecting membrane can be regarded as a continuous sheet.

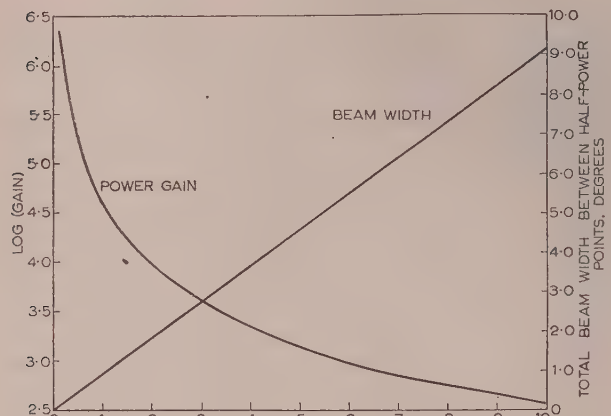


Fig. 1.—Calculated power gain and beam width of the 250 ft aperture radio telescope as a function of wavelength.

The calculations have assumed a plano-sinusoidal distribution of intensity across the aperture. The power gain is relative to an isotropic radiator, and an efficiency of 60% has been assumed.

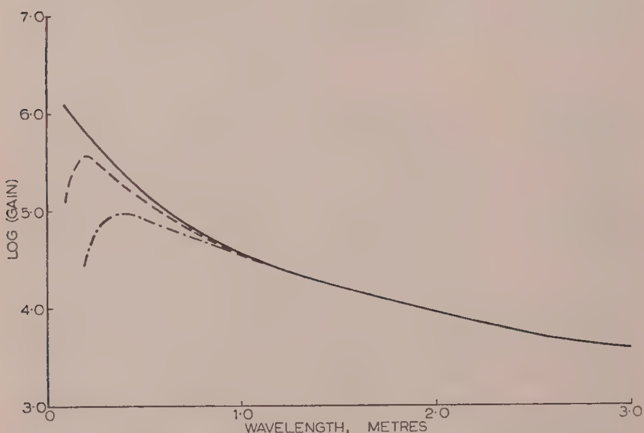


Fig. 2.—Effect of distortion on the power gain of the 250 ft telescope at short wavelengths.

The calculations have assumed irregularities of ± 2 in (—·—·), ± 1 in (— — —) and $\pm\frac{1}{2}$ in (————), with a Gaussian distribution relative to the focus.

(2.2) Beam Width and Power Gain

The results of theoretical calculations of the power gain and beam width of the telescope are shown in Figs. 1 and 2.* The beam width calculations have been made in the conventional manner for a circular paraboloid assuming a plano-sinusoidal distribution across the aperture. The figures are for the total width of the beam between half-power points. The power gain is expressed relative to an isotropic radiator, and an efficiency of 60% has been assumed. At the shorter wavelengths the distortions and irregularities in the bowl will begin to influence the power gain significantly. The extent of these departures from the true parabolic shape will depend on many factors, particularly the wind, and cannot be assessed finally until the telescope is brought into use. It is possible, however, that under good conditions the shape will remain true to within $\pm\frac{1}{2}$ in, but ± 1 in may be a more realistic figure for general operational conditions. Fig. 2 shows the effect of such distortions on the power gain calculated on the assumption that the irregularities have a Gaussian distribution relative to the focus.

The data of Figs. 1 and 2 will be checked experimentally as soon as the telescope is in an operational condition by observation of the radio sources on various wavelengths.

(2.3) The Motion of the Telescope

The elevation and azimuth movements of the telescope are independently controlled by electric motors in a Ward Leonard system. In practice the basic elevation and azimuth motions will rarely be used under observational conditions, and it has been necessary to design a complex control system in order to give the sidereal, scanning and other automatic motions which will be required.

The problem of sidereal motion with an alt-azimuth mounted instrument can be solved by means of a mechanical analogue, but this has a number of restrictions (such as a difficulty in following a star past the zenith) which it was desired to avoid. Dr. J. G. Davies, of Jodrell Bank, therefore developed an alternative electrical method. This uses magslip resolvers to give signals proportional to the sine and cosine of an angle, in servo loops to solve the fundamental equations of spherical trigonometry. The range of automatic following speeds required varies from about 10° per hour for a star near the southern horizon to infinity for a star passing through the zenith. In practice the maximum speed will be between 20° and 30° per minute, and this will be adequate except within a very small zenithal angle. In addition to this fundamental sidereal motion, the telescope will be required to follow the sun, moon and the planets, and additional parallax corrections can be switched in when needed.

The control system enables the telescope to be positioned at any predetermined azimuth-elevation, right ascension-declination or galactic latitude-longitude, and to maintain these positions or alternatively to perform an automatic scan at a rate which can be varied between 2° per hour to 5° per minute. The separation of the successive scanning lines can be varied between ½ and 4° to suit the beam width.

No slip rings are used on the telescope, and hence the motion in azimuth is limited to a little over 360°. However, occasions will arise when it will be necessary to follow a source through the positions of the stops, and arrangements are made in the control system for this reversal to take place automatically.

The tests of the prototype control system proved very satisfactory and gave a following accuracy of ± 5 min of arc. The final following accuracy of the telescope will probably be determined by the mechanical system. The original specification called for an accuracy of ± 12 min of arc to be maintained up to speeds of 4° per minute both in azimuth and elevation.

(2.4) The Construction of the Telescope

Fig. 3 is an artist's impression of the engineer's design for the radio telescope based on the above specifications. The main paraboloidal bowl is carried on the steel towers at a height of

* I am indebted to Dr. C. Hazard of Jodrell Bank for performing these calculations.

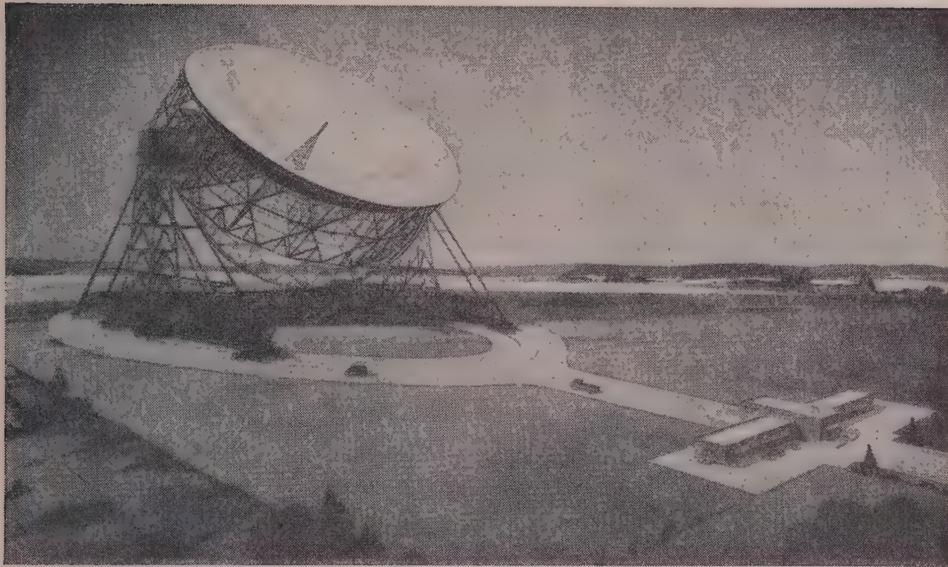


Fig. 3.—The artist's drawing of the 250 ft aperture steerable telescope.

ıbout 170 ft above the ground. Motion in elevation will be ıchieved by four 25 h.p. electric motors, situated in the labora- ories at the top of these towers, and driving through 25 ft racks aken from the battleship *Royal Sovereign*. The towers are each carried on bogies, the inner two of which are driven, the outer ones serving as wind carriages. These run on the 17 ft- gauge double railway track, again driven by four 25 h.p. motors. The cross-girder which connects the base of the two towers carries the telescope on a central pivot which is its fundamental ocating point. Both this pivot and the 350 ft-diameter railway rack are supported on deep piled foundations which extend rom 45 ft to 90 ft underground. Some 10 000 tons of reinforced concrete have been used in these foundations in order to give he necessary stability to the instrument. The total weight of the steel superstructure is about 2 000 tons, of which 700 tons are concentrated in the bowl framework. The weight of the steel membrane with which the bowl is surfaced is 70 tons.

(3) THE INVESTIGATION OF RADIO WAVES FROM THE GALAXY

The pioneer observations of Jansky[3,4] and later of Reber[5,6] showed that the intensity of the radio emission varied markedly with the direction of the aerial beam, being most intense from he direction of the galactic centre. In fact the variation was generally that which might be expected if the stars in the milky way system were responsible for the emission. However, neither Jansky nor Reber, nor any subsequent worker, has succeeded in detecting radio emissions from any of the stars (other than he sun); neither have the localized radio sources since identified coincided with any typical common star. It is possible that Reber's original suggestion, that the radiation is emitted by the interstellar gas, is at least partially true, but the situation is most complex, and the question of the origin of these galactic radio emissions will form a prominent part of the programme of the new telescope.

When it was discovered that some of the radiation from space was emanating from localized radio sources, it was tempting to conclude that the background continuum was made up of large numbers of these discrete radio sources, unresolved by the available radio techniques. This situation would be analogous to that in which the milky way is viewed by eye or through a telescope of low resolving power, when all the faint stars appear as a continuum of light and only the brightest stars stand out individually. It is now known that this view of the radio emission is untenable; not only have improved techniques failed to reveal the increased number of appropriately distributed sources, but the spectra of the continuum and the sources are not compatible. It is therefore convenient to review the present situation in this light.

There is a fundamental difficulty in studying the distribution of the radio emission because of the appreciable width of the beam inevitable in the surveys with existing radio telescopes. However, some recent distributions obtained by Mills[7] with the high-definition 'cross' type of aerial at Sydney show the nature of the problem very clearly. Some sections across the galactic plane near the galactic centre are shown in Fig. 4. It appears that distributions of the following types are involved:

(*a*) One very sharply concentrated towards the galactic plane with a width between half-brightness points of about 3°. This component

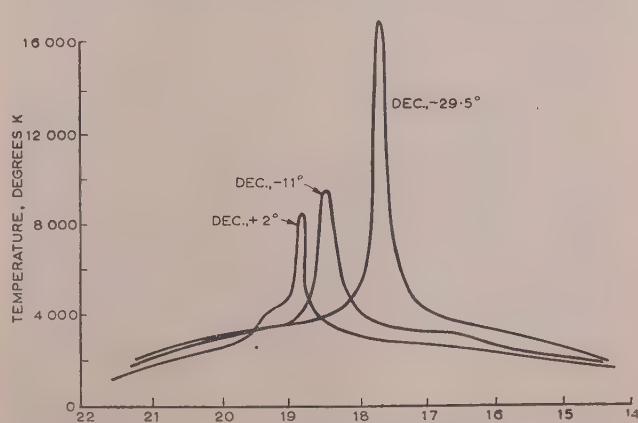


Fig. 4.—Some examples of the distribution of radio intensity across the galactic plane near the galactic centre.

Obtained with a high-definition aerial by Mills on a wavelength of 3·5 m.

is also markedly concentrated towards the galactic centre, although less sharply than across the plane, and appears to be distributed in a similar manner to the interstellar gas.

(*b*) A broad and flat component with an estimated angle between half-brightness points of about 60°. This forms a 'corona' to the Galaxy, with no obvious parallel in galactic structure.

(*c*) Some workers[8,9,10] believe that the surveys cannot be interpreted in terms of (*a*) and (*b*) but that a third, intermediate, component (*c*) is required and best exemplified in Fig. 4 by the knee between the (*a*) and (*b*) distributions in the section at declination +2°. This component has a distribution similar to that of the mass of the Galaxy.

(*d*) Finally, the surveys of localized sources have revealed a concentration of the most intense and extended sources in the plane of the Galaxy. This component forms part of the sharply concentrated component (*a*), but as mentioned above is certainly distinct from it in origin.

Some possible suggestions as to the mechanisms by which these components arise are as follows:

(*a*) It is almost certain that a significant fraction of the component concentrated in the galactic plane arises as a thermal emission in the ionized gas (i.e. in H II regions), by the process of free-free transitions originally suggested by Reber. Such thermally emitting gas will have a temperature spectrum given by $T \propto \lambda^n$, where n varies between 0 and 2 depending on the opacity. When the opacity is small the observed intensity will be independent of wavelength, and it is highly probable that at short wavelengths the component (*a*) is predominantly due to this mechanism. On the other hand, it is impossible to fit the observations at the longer wavelengths on this assumption where a non-thermal component must predominate. The origin of this non-thermal component is quite unknown; speculations are along the lines that it may arise through the acceleration of high-energy electrons in a turbulent gas. The relativistic electrons would then radiate at radio frequencies in the associated magnetic fields.

(*b*) The main suggestion about the origin of the broad distribution is due to Shklovskii.[11] He concludes that the distribution is roughly spherical, concentric with the galactic centre, and that the emission is of a synchrotron type from relativistic electrons moving in weak magnetic fields. This distribution can be regarded as a 'corona' around the Galaxy with a radio emission which has a non-thermal spectrum. Whereas the width of the (*a*) component of about 400 parsecs ($\equiv 3°$) is comparable with the thickness of the neutral hydrogen gas in the galactic plane, the radius of this corona must be of the order of 10 kiloparsecs and does not correspond with any known feature of galactic structure.

(*c*) As indicated above, certain groups believe that a third component of the continuum exists with a distribution similar to that of the galactic mass. However, there are no valid suggestions as to its origin. The common stars are certainly not responsible, and it seems that the radio stars so far discovered (see below) are too few to contribute appreciably to this component.

(*d*) For some time after the discovery of the first few dozen localized sources it was believed that these were mostly in the Galaxy. The recent surveys, which will be considered later, have shown that this view is untenable and that the majority of the sources are probably extragalactic. However, there is a concentration of the intense sources in the galactic plane. For example, in the Cambridge survey[12] of 1936 sources there were 30 extended sources (i.e. with diameters greater than 20 min of arc), and half of these with intensities greater than 10^{-24} watt/m^2 per c/s were concentrated within about 15° of the galactic plane. In the southern hemisphere, Bolton, Stanley and Slee[13] found that of 30 sources with intensities greater than this, half were within 5° of the galactic plane. In a few cases either the distance has been measured or an identification with a visible object has been made, and it is safe to assume that these sources are in the Galaxy and form part of the sharply concentrated component (*a*). It cannot be assumed that there is necessarily any connection between the concentrated continuum and the localized sources other than a spatial association. The nature of these localized sources is discussed in the next Section.

(4) THE NATURE OF THE LOCALIZED SOURCES IN THE GALAXY

Between 2000 and 3000 localized sources of radio emission have so far been found in the surveys of both hemispheres. Of these about 15 or 20 satisfy the various criteria, such as appreciable angular extent, intensity and distribution, which make it highly probable that they are members of the local Galaxy. Seven of these have been satisfactorily identified with galactic objects. These are so unusual that they will be discussed separately in this Section. Surveys at centimetre wavelengths show additional localized sources, but these are associated with well-known ionized H II regions and are excluded from present considerations.

(4.1) Remnants of Supernovae

Shortly after the discovery of the existence of localized radio sources, Bolton, Stanley and Slee[14] suggested that one of the intense sources might be associated with the Crab nebula. Subsequent more accurate measurements of the position of the radio source and of its angular dimensions leave no room for doubt that the radio source in Taurus, which is the third most intense in the sky, is in fact associated with this nebula. The Crab nebula is the expanding gaseous shell of the supernova of A.D. 1054. Its distance is about 4000 light-years, and the angular dimensions of both the telescopic and radio object are about 4×6 min of arc. The temperature of the gaseous shell, which is expanding at the rate of about 70 million miles per day, is 50000° K, which is far too low to produce the observed radio intensity of $1{\cdot}8 \times 10^{-23}$ watt/m^2 per c/s (at 80 Mc/s) by thermal processes. Interest in this object has been intensified since the discovery in Russia[15] that the light from the nebula is polarized, and the development of theories by Shklovskii[16] and subsequently by Oort and Walraven[17] that the polarization and the radio emission could be explained as a synchrotron mechanism resulting from the movement of high-energy electrons in weak magnetic fields. It now seems highly probable that supernovae like the Crab nebula are both powerful radio emitters and responsible for the generation of an appreciable fraction of cosmic rays.

There are two other well-attested cases of supernovae in the Galaxy, those observed by Tycho Brahe in 1572 and by Kepler in 1604. Unlike the Crab nebula, these are not spectacular objects; in fact, until the recent work of Baade, no remnants of Tycho's object had been located. Even so there seems little doubt that radio sources are associated with these supernovae. The identification of a radio source with the 1572 supernova was made by Hanbury Brown and Hazard[18] in 1952, and with the 1604 supernova by the Cambridge group[12] in 1955.

(4.2) The Cassiopeia Radio Source

The most intense radio source in the sky lies in Cassiopeia. Its angular extent is about 4 min of arc, and the flux density at 80 Mc/s is $2{\cdot}3 \times 10^{-22}$ watt/m^2 per c/s. Although this was the first radio star to be discovered in the northern hemisphere, by Ryle and Smith,[19] it was not until 1951 that a successful search for the visible counterpart was initiated with the 200 in telescope by Baade and Minkowski.[20] The result of this search was the surprising discovery that this powerful radio emitter consists of a very faint extended nebulosity of a type previously unknown. The gaseous filaments in this nebulosity are in violent motion at thousands of kilometres per second. There is no satisfactory explanation of the mechanism of the generation of radio waves; neither is there any agreement as to the nature of the object itself, although opinions have been expressed that it may be the remains of a very old supernova.

(4.3) Other Identifications

There are only three other agreed identifications of radio sources with galactic objects. These are the Cygnus loop[21] and the nebulosities in Auriga[22] and Gemini[23] (I.C.443). These are all extended gaseous nebulosities of low photographic brightness, with filamentary structure. As with the Cassiopeia source, there is no agreed opinion as to the nature of these objects or of the mechanism whereby they emit radio waves.

The only general conclusion which can be drawn from the present situation is that the galactic radio sources appear to represent a very rare type of celestial object characterized by an appreciable extension of diffuse gas of low photographic brightness. Whether they are different manifestations of the same phenomena (e.g. supernovae) or vary in character remains uncertain. Attempts have been made to associate the radio sources with other rare classes of galactic objects such as novae, planetary nebulae and globular clusters, but without success.

The new telescope, with its high definition and adaptability over a wide frequency range, is expected to be a powerful tool in the investigation of the problem of these galactic sources and of the continuum. Initially it is hoped to make measurements at a few selected points in the frequency range from about 20 Mc/s to 1 400 Mc/s in order to study the isophotes of the continuum and the spectra of the localized sources.

(5) THE INVESTIGATION OF RADIO WAVES FROM EXTRAGALACTIC NEBULAE

For a considerable time there was uncertainty whether the local Galaxy was unique in its production of radio waves or whether the extragalactic nebulae were similarly endowed. After several inconclusive results, Hanbury Brown and Hazard[24] in 1950, using the 218 ft transit telescope at Jodrell Bank, were able to show that the great spiral nebula in Andromeda (M31) was a radio source and that the total radiation from this galaxy was of the same order as the integrated emission from the local Galaxy. Shortly afterwards, Baade and Minkowski[20] announced that they had identified the second most intense radio source in the sky in Cygnus, with two galaxies in collision at a distance of 200 million light-years. Subsequent measurements of the angular diameter and shape of the radio source, and more recently of the hydrogen line 'red shift' (see Section 7) confirm this identification. The existence of this intense radio source ($1{\cdot}4 \times 10^{-22}$ watt/m^2 per c/s at 80 Mc/s) associated with a celestial collision nearly at the limit of penetration of the 200 in telescope is one of the most remarkable features of radio astronomy with far-reaching implications which will be discussed in a moment.

At the present time the surveys in England and Australia have revealed between 2 000 and 3 000 localized sources. Apart from the galactic concentration of a small number of the intense and extended sources discussed in Section 4, these sources are distributed isotropically and are probably extragalactic. A relatively small number have been identified with telescopic objects. We shall review these first and reserve the discussion of the large number of unidentified extragalactic sources for the next Section.

(5.1) Normal Extragalactic Nebulae

After the measurements on the Andromeda nebula M31, Hanbury Brown and Hazard[25] succeeded in measuring the emission from six further extragalactic nebulae which were in the field of view of the transit telescope. More associations have been made as a result of the Cambridge survey,[12] and a further thirteen have been studied in the southern hemisphere, including the Magellanic Clouds, by Mills.[7] The relation between the radio and photographic magnitudes of these normal nebulae so far studied is shown in Fig. 5. If the galaxies had a constant ratio of radio to optical emission the points would lie on a line with slope of unity. The line of best fit plotted in Fig. 5 corresponds to $(m_R - m_p)$ of $+1{\cdot}4$. The galaxies so far studied consist only of types Sb, Sc and the Magellanic Clouds, and the data are too scarce to comment with much realism on the variation of radio emission with type of nebulae. The first important problem to be settled is the method by which these

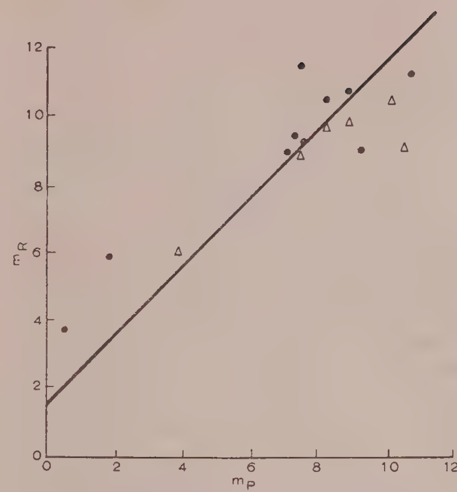


Fig. 5.—A comparison compiled by Mills of the radio (m_R) and optical (m_P) emission of all normal galaxies whose radio emission has been detected.

● Southern galaxies.
△ Northern galaxies.

The line of best fit corresponds to $(m_R - m_P) = +1{\cdot}4$.

normal galaxies generate the radio waves, whether different types of galaxies vary in their emission, and whether such variations are associated with differences in stellar population and the place of the galaxy in the evolutionary sequence. These problems are, of course, closely related to the problem of the radio emission of the local Galaxy discussed in Sections 3 and 4. In the case of M31, evidence[26] has already accumulated that a component of the emission may be in the form of a 'corona' similar to one of the components of the continuum of the galaxy.

(5.2) Abnormal Extragalactic Objects

The association of the second most intense radio source with the collision of two spiral nebulae has been mentioned above. During the last few years a further half dozen or so radio sources associated with unusual extragalactic objects have been discovered. These include NGC1275 in the Perseus cluster, and a source in Hydra, which Baade and Minkowski consider may be galaxies in collision. Other peculiar associations include NGC5128, which has a dark band across it; and M87, from which a jet emanates.

The normal nebulae show a ratio of radio to optical emission of about 10^{-6}. Compared with this, the peculiar objects have a much greater ratio, of the order of 10^{-3}, reaching unity in the case of Cygnus. As pointed out by Pawsey and Bracewell,[27] this represents an extremely high conversion efficiency, and the mechanism by which such objects generate radio waves is a challenging problem. In a collision of galaxies the stars are too widely separated for significant collisions to occur, but the dust and gas, which represent an appreciable fraction of the galactic mass, will certainly suffer real collisions at velocities of, perhaps, thousands of kilometres per second. The key to the mechanism of generation of the radio waves probably lies in this highly agitated ionized gas.

The collection of further data on both the normal and abnormal radio sources is essential to an understanding of the problem of the extragalactic radio emissions. The new telescope is well fitted to pursue this important task over a wide range of wavelengths.

(6) THE UNIDENTIFIED EXTRAGALACTIC SOURCES AND THE COSMOLOGICAL PROBLEM

A common feature of the surveys of radio stars in both northern and southern hemispheres is the failure to associate all but a few with celestial objects visible in the large telescopes. For example, the Cambridge survey located 1 936 radio stars, 1 906 of which were of small angular diameter and distributed isotropically. Only a few per cent of these have been successfully associated with the known objects discussed in the previous Sections, and the surveys in the southern hemisphere by the Sydney group give similar results.

The spatial distribution of these unidentified radio stars has been investigated by plotting log N against log I for each of a number of areas of sky, where N is the number of radio sources per unit solid angle with an intensity greater than I. If the sources are distributed uniformly throughout space it is easy to show that the plot should give a straight line of slope $-1 \cdot 5$, so that any departure from this relationship will give information about the variation of spatial density with distance. The method was first used by Mills[28] and by Bolton, Stanley and Slee[29] to investigate the distribution of a few hundred sources located in the earlier surveys, but here we shall refer only to the later analysis by Ryle and Scheuer[30] of the 1 906 sources in the Cambridge survey and to the preliminary results of a similar southern hemisphere survey by Mills[31] using the large 'cross' aerial.

The log N/log I Cambridge plots for seven areas of sky are shown in Fig. 6, where the straight line represents a slope of $-1 \cdot 5$, such as would occur with a uniform spatial density. The results are very surprising. The seven areas are similar, but all show a marked departure from the expected slope, the experimental curve becoming steeper for low intensities until a flattening due to instrumental limitations occurs. The implication is that the spatial density of radio stars is constant in the neighbourhood of the solar system but progressively increases with distance. Ryle and Scheuer show that it is impossible to account for this result and for the isotropy of the sources other than by assuming that the increase in slope is taking place at distances which are comparable with the limits of observation of the 200 in telescope. If, in fact, this spatial distribution is due to processes on a cosmical scale, then the results have a significant bearing on cosmological theories. In the steady-state or continuous-creation theories the density of nebulae should be everywhere the same and independent of time and place, since new galaxies are continually being formed to take the place of those which move out of the field of view as a result of the expansion of the universe. On the other hand, in the evolutionary theories the spatial density decreases progressively with time. In this case we should expect to find a greater concentration of nebulae near the limits of the observable universe which corresponds to a period of time some 2 000 million years ago.

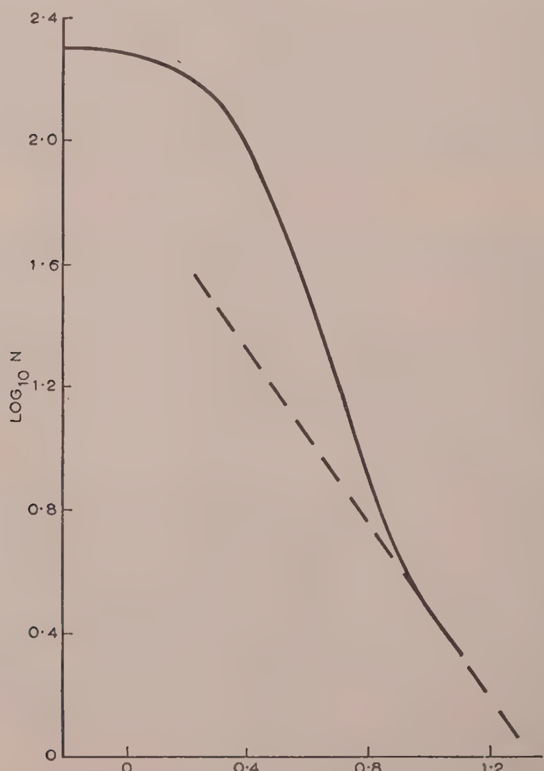


Fig. 6.—Number/intensity distribution on a logarithmic plot of the radio sources detected in the Cambridge survey.

The straight dashed line corresponds to the slope of $-1 \cdot 5$ which would apply if the sources had a uniform spatial density.

This is precisely the result which appears in the distribution of Fig. 6, but since the radio emission from individual nebulae is too weak to be detected at such distances we are led to further conclusions about the nature of these radio sources. Galactic collisions of the Cygnus type can be shown to produce radio emission of the order of magnitude required, and the suggestion is that these sources are colliding galaxies probably lying beyond the limit of penetration of the 200 in telescope. This is consistent with the interpretation of the log N/log I curve, since at the time of several thousand million years ago both the spatial density and the number of colliding galaxies must have been much greater than at present.

This interpretation of the log N/log I curve for the observed radio sources is therefore in favour of the evolutionary cosmological theories. This conclusion is, at present, in very great dispute, particularly as the preliminary results of the Sydney survey do not entirely agree with those presented by Ryle and Scheuer. The issue is of the utmost importance, carrying with it the possibility of a significant contribution to the main cosmological problem of the origin of the universe. Before agreed conclusions can be reached the validity of the data must be greatly increased, and this programme can be expected to form a very prominent aspect of the work of the new telescope. In the programme as planned now, the telescope will be used as an element in an interferometer to measure the actual angular extent of the sources so that the effective temperatures of the individual sources can be determined. If, in addition, the expectation that the distances can be measured, as described below, is fulfilled, then precise data will be available about the spatial distribution of the different classes of extragalactic radio sources.

(7) THE SPECTRAL LINE EMISSION FROM NEUTRAL HYDROGEN GAS

In the hydrogen atom in the ground state there are two possible orientations of the electron spin in the weak magnetic field associated with the nuclear magnetic moment. The energy difference between these two states is such that an electron transition will give rise to emission at a frequency of 1 420·405 Mc/s. The probability of the spontaneous emission of such photons is very low—a hydrogen atom in free space would undergo one such transition in 11 million years on the average. Even so, and in spite of the fact that the average density of hydrogen in interstellar space is only about 1 atom per cm^3, van de Hulst[32] predicted in 1945 that such radiation should be detectable. His

prediction was brilliantly verified in 1951, almost simultaneously at Harvard,[33] and in Holland[34] and thereafter in Australia.[35]

This discovery opened a new era in radio astronomy, because for the first time it became possible to study the clouds of neutral hydrogen gas in the Galaxy, and, through a measurement of the Doppler shift in frequency of the line, to determine the velocity of the clouds relative to the earth. The pursuit of this avenue of galactic exploration by the astronomers at Leiden during the past few years is particularly outstanding. Because of the dense optical obscuration of the nuclear region, very little detail was known about the structure of the Galaxy. The possibility of studying the hydrogen line emission removed these hindrances, and most remarkable information is now available about the spiral structure of the galactic system.[36]

The purely astronomical implications of this work were immediately apparent, and at least three observatories put in hand the construction of large radio telescopes for the specific purpose of investigating further this 21 cm radiation from galactic hydrogen. Telescopes of 83 ft diameter were opened in March, 1956, at Dwingeloo (Holland) and in September at Bonn, and one of 60ft diameter was opened in April at Harvard. They will enable even more precision to be obtained in the research on galactic structure, and should also lead to valuable preliminary information about the structure of the Andromeda nebula M31.

As will be evident from Section 2, there are grounds for hoping that under good weather conditions the Jodrell Bank telescope will be capable of working on 21 cm, and preparations for its use on this wavelength are in hand. There is no present intention to compete in the field of galactic structure which is so admirably pursued at the observatories mentioned above. Instead it is intended to develop the absorption techniques for the measurement of the distances of the galactic radio sources already initiated with smaller telescopes at Jodrell Bank by Williams and Davies,[37] and for measurements of the 'red shift' of the extragalactic sources.

The principle of the absorption method is illustrated in Fig. 7, which shows some measurements on the Cassiopeia radio source. It is assumed that the emission is continuous and does not show line structure at 21 cm. When the radiation traverses the neutral

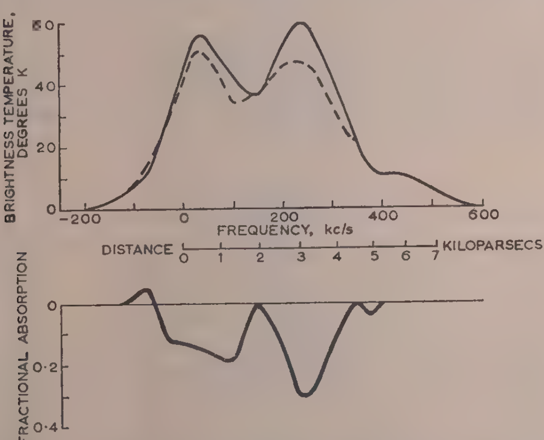


Fig. 7.—An example of the absorption method for measuring the distance of radio sources.

——— Spectrum expected in absence of source.
– – – – Spectrum observed.
▬▬▬ Absorption spectrum.

The zero of frequency on the abscissa corresponds to the hydrogen line frequency of 1 420 Mc/s unaffected by the Doppler shift which displaces the frequency at which the absorption occurs. There are two peaks in the spectrum because in this case the radio emission from Cassiopeia traverses two spiral arms with different velocities relative to the earth. The distance equivalent to the frequency shift is shown below the abscissa.

hydrogen in the Galaxy there will be absorption at the line frequency, whereas the emission at neighbouring frequencies will be unaffected. Hence, if the spatial distribution of the hydrogen is known, the distance of the source can be computed from the extent of the absorption. The absorption will not necessarily occur at 21 cm, since, in general, there will be a Doppler shift due to the motion of the hydrogen clouds.

This particular method can be used only for sources situated within the Galaxy. The only information obtained about extragalactic sources is that they are, in fact, extragalactic. It is evident from Section 6 that the determination of the distances of the extragalactic sources is of primary importance in the investigation of the major cosmological problem, and it is in this sphere that the high discrimination and gain of the new telescope will be particularly valuable. In the normal telescopic measurements of distance of extragalactic nebulae, the red shift of known spectral lines is determined, and the distance is then obtained from Hubble's relation between the distance and the velocity of recession. The proposed radio method is analogous, but the technical difficulties are severe because of the extent of the shift of line frequency. For example, the speed of recession of the Cygnus source is 16 800 km/s and the corresponding frequency shift of the 21 cm line is 80 Mc/s. Thus, investigations of this type demand the development of apparatus capable of sweeping the frequency range in a continuous band from below 1 300 to 1 420 Mc/s. The small 24 ft radio telescope at Harvard has recently demonstrated the potentialities of this method by determining the red shift on the hydrogen line emission from the Coma cluster of galaxies. Heeschen[38] measured the line in emission from this cluster at a frequency of 1 386 Mc/s, corresponding to a recession velocity of 7 000 km/s, in very reasonable agreement with the optically determined red-shift velocity of 6 680 km/s. In order to make this measurement, Heeschen had to integrate 80 drift curves across the cluster, and the technique is rather limited and inaccurate when used with such small telescopes. Calculations indicate that in the case of clusters the Jodrell Bank telescope should give a clearly measurable deflection on a single-drift curve up to shifts of about 20 Mc/s, corresponding to recession velocities of about 4 000 km/s and distances of the order of 70–80 million light-years. Integration of numerous records with this instrument would probably enable the distances of clusters to be measured near to the limit of possible exploration of the universe.

There is fortunately a further possibility, already demonstrated by Lilley and McClain[39] with the 50 ft radio telescope at the Naval Research Laboratory in Washington, which will probably ease the problem of measuring the red shifts of these radio sources at great distances. Success in this method depends on the possibility that in colliding galaxies the intense continuum radiation may have to pass through neutral hydrogen gas associated with the colliding galaxies before it passes into intergalactic space. In this case the radiation in the continuum at the spectral line frequency will be absorbed in this gas and may then be observable in absorption. The frequency at which the absorption is observed will, of course, be that appropriate to the frequency shift of the line associated with the recession. Although at the limit of their technique, Lilley and McClain have successfully observed this absorption in the remote Cygnus source by integrating a large number of records. They found an absorption band in the continuum centred at about 1 340 Mc/s, corresponding to a recessional velocity of 16 700 km/s, in very good agreement with the optical red-shift measurements.

Hence there are good grounds for anticipating that the new telescope will be able to measure the distances of large numbers of radio sources, both galactic and extragalactic, by means of these 21 cm spectral line observations.

(8) THE TELESCOPE AS A TRANSMITTER: LUNAR AND PLANETARY INVESTIGATIONS

The previous examples of the use of the new telescope have all concerned the programmes in which the instrument will be used as a receiving aerial. The original proposals included many problems in which it would be used as a combined transmitting and receiving aerial. In these radar or radio-echo aspects of the work, the moon and the planets will figure prominently.

Radio echoes from the moon were first claimed to have been observed in 1946 by Z. Bay[40] in Hungary. His recording system was unusual, and the first certain echoes to be obtained in the conventional sense on a cathode-ray tube were by the U.S. Army Signal Corps.[41] Subsequently echoes were obtained by Kerr and Shain[42] in Australia. These experiments showed that the moon echoes were subject to deep and rapid fading—an effect which is now believed to be due to a peculiarity of the moon's motion with respect to the earth, known as libration. During this period, apparatus for lunar echo studies was also under development at Jodrell Bank, and it appears that this is the only systematic investigation of the moon by the radio echo technique which has yet been carried through.[43] This apparatus works on a frequency of 120 Mc/s and uses a transmitter giving 10 kW in 30 millisec pulses at a recurrence rate of 0·6 per second. The receiver bandwidth is 30 c/s, and appropriate arrangements have to be made to allow for the Doppler shift in the frequency of the returned signal. The most important results obtained with this apparatus concern the long-period fading (20 to 30 min), which by cross-polarization experiments[44] has been shown to be due to the rotation of the plane of polarization of the radio wave as it traverses the ionosphere (the Faraday effect). This immediately led to the development of a moon echo system by which the *total* electron content of the ionosphere could be determined.

The magneto-ionic theory shows that the polarization shift Ω is given by

$$\Omega = \frac{k}{f^2}\int N(r)dr \text{ complete rotations}$$

where k is a factor including the geomagnetic field and the inclination of the geomagnetic vector to the line of sight; f is the radio frequency; $N(r)$, the electron density; and dr, the element of path length. If the rotation occurs in a radio wave which has penetrated the entire ionosphere, as in the case of the moon echo, then $\int N(r)dr$ is the total electron content of the ionosphere per square centimetre along the line of sight. Experiments on a single frequency can only give the rate of change of $\int N(r)dr$, and in order to evaluate the integral it is necessary to carry out the measurements on two closely spaced radio frequencies. Measurements of the total electron content by this means have been made by Evans,[45] and an example of his preliminary results is shown in Figs. 8 and 9. Fig. 8 shows the

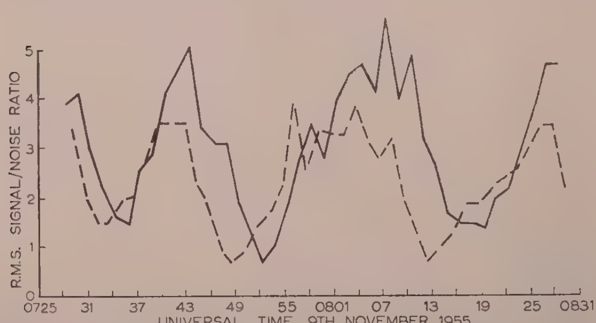


Fig. 8.—R.M.S. signal/noise ratio of lunar echoes.
——— 120·72 Mc/s.
— — — 119·28 Mc/s.

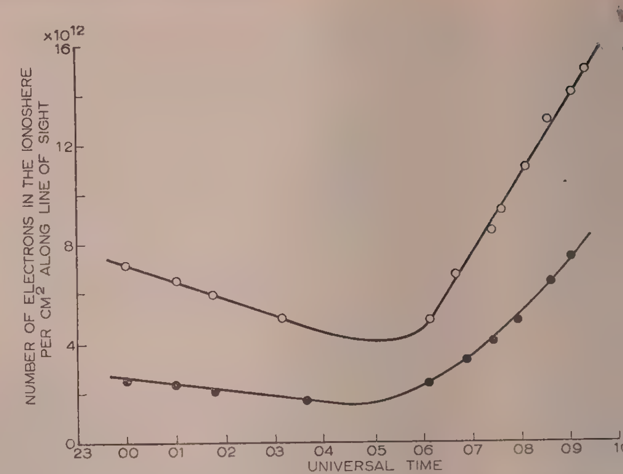


Fig. 9.—Total ionospheric electron content per square centimetre along the line of sight for October, 1955.
○—○—○ Obtained from lunar echo data of type shown in Fig. 8.
●—●—● Calculated from Slough critical-frequency data assuming parabolic distribution of density with height.

r.m.s. signal/noise ratio of the lunar echoes at frequencies of 120·72 and 119·28 Mc/s. The variation is similar but there is a time displacement between the two curves. Fig. 9 shows the total electron content deduced from such measurements, together with that calculated from critical frequency data assuming a parabolic distribution. These results indicate that the electron content is considerably greater than that estimated on the basis of a simple parabolic region.

The technical difficulties in this work are considerable, and with the present aerial system, measurements can be made only with the moon in transit for about 10 periods in each lunation. The new telescope will immediately remove these handicaps and will enable systematic data to be collected about the total ionospheric electron content. This is bound to be of considerable importance to our understanding of the ionosphere and of solar-terrestrial relationships.

The problem of radio echoes from the planets is vastly more difficult, and, as far as is known at present, no serious attempts have yet been made to solve it. The magnitude of the problem relative to the moon is indicated in Table 1, which gives the distance d, the radius r and a factor $\pi r^2/d^4$, which is the appropriate variable in the radar equation.

Table 1

Target	Distance from earth, d	Radius, r	$\pi r^2/d^4$	Value of $\pi r^2/d^4$ relative to the moon
	km	km	m^{-2}	
Moon ..	$3{\cdot}84 \times 10^5$	$1{\cdot}74 \times 10^3$	3×10^{-22}	1
Venus ..	$3{\cdot}7 \times 10^7$	$6{\cdot}1 \times 10^3$	6×10^{-29}	2×10^{-7}
Mars ..	$5{\cdot}3 \times 10^7$	$3{\cdot}4 \times 10^3$	4×10^{-30}	$1{\cdot}3 \times 10^{-8}$
Mercury ..	$7{\cdot}7 \times 10^7$	$2{\cdot}4 \times 10^3$	5×10^{-31}	$1{\cdot}7 \times 10^{-9}$
Jupiter ..	$5{\cdot}8 \times 10^8$	$7{\cdot}1 \times 10^4$	1×10^{-31}	$3{\cdot}3 \times 10^{-10}$
Saturn ..	$1{\cdot}2 \times 10^9$	$6{\cdot}0 \times 10^4$	5×10^{-33}	$1{\cdot}7 \times 10^{-11}$
Uranus ..	$2{\cdot}5 \times 10^9$	$2{\cdot}5 \times 10^4$	5×10^{-35}	$1{\cdot}7 \times 10^{-13}$
Neptune ..	$4{\cdot}2 \times 10^9$	$2{\cdot}6 \times 10^4$	7×10^{-36}	$2{\cdot}3 \times 10^{-14}$
Sun ..	$1{\cdot}5 \times 10^8$	$7{\cdot}0 \times 10^5$	3×10^{-27}	1×10^{-5}

The last column shows that success in detecting radio echoes from Venus would demand an overall power sensitivity between a million and ten million times greater than that required in the case of the moon. This assumes, of course, that the reflection

coefficient of the planet would not be inferior to that of the moon. The problem cannot be appreciably eased by increasing he length of the transmitter pulse with appropriate decrease of eceiver bandwidth, because of the Doppler spread introduced by he rotation of the planet. The rotation period of Venus is nknown (this would, in fact, be one of the main scientific results o be expected from the experiment), but on the basis of current stimates of the rotation period, the Doppler spread would robably limit the useful pulse width to about 40 millisec, which s only a few times longer than that used in the lunar investiga-ions. The main factor must therefore be achieved in the gain f the aerial, by increasing the transmitter power, and possibly y integration of successive echoes.

The problem has been carefully considered at Jodrell Bank in elation to the very great gain of the new telescope, and an ttempt to obtain planetary echoes will be made early in the esearch schedule. The complete return journey of the earth–Venus radio signal will take 4 min, and success in detecting such radio echo would be a spectacular technical accomplishment. Nevertheless, the experiment could not be justified on this basis, nd it is hoped that with the telescope a systematic programme vill be possible, in which the rotation period can be determined nd information obtained about the Venusian surface and tmosphere.

A further interest in this planetary work is the possibility of neasuring the range of the planet with sufficient accuracy to nprove our knowledge of the solar parallax. The present osition is that the two most accurate optical measurements hould be correct to 1 part in 10 000 but differ between themselves y 1 part in 1 000. The possibility of improving this measure-nent by the radio echo technique will certainly receive immediate ttention if the initial attempts at detection are successful.

On the basis of Table 1 it should require about 100 times less ower sensitivity to obtain a radio echo from the sun than from Venus, but much of this gain would be lost because of the radio oise emitted by the sun, which will degrade considerably the ffective receiver sensitivity. There are also major uncertainties onnected with the manner in which an incident radio wave will e reflected in the sun's gaseous atmosphere which make com-arison with the moon and planets very unreliable.

If the planetary experiment is successful, another major rogramme of work in radio astronomy will be opened. For xample, many of the larger and more closely approaching steroids will come within the scope of the telescope, and some spects of the intriguing problem of the origin of the *Gegenschein*, r counterglow, could be studied.*

(9) SOME PROBLEMS IN METEOR ASTRONOMY AND PHYSICS

Considerable progress has been made during the last ten years s a result of the application of radio echo techniques to the study f meteors.† The ability to measure velocities and radiants of he meteors by studying the radio echo scattered from the ionized rail under any sky conditions has given systematic information bout the distribution of sporadic meteors and of those con-entrated in showers. In this work the meteors in the visual ange down to magnitude +5 or +6 are readily covered by the ensitivity of quite small equipments, and the range has been xtended to magnitude +10 without much difficulty. A clear roof has been given that all the sporadic meteors are members f the solar system moving in short-period orbits around the un, and a vigorous series of meteor showers incident on the sunlit hemisphere has been discovered. Some of these showers are clearly related to comets; for others there is no obvious parent body, and, like that of the sporadic meteors, their origin is at present an enigma. The number and mass of these meteors entering the earth's atmosphere are very considerable—several thousand million, bringing in a mass of between 1 000 and 2 000 kg, per day. The primary problem in meteor astronomy is to unravel the connection between this debris and the solar system—is it some of the primeval matter, or has it arisen from collisions and disintegrations of planetary or asteroidal bodies at a later date?

The new telescope will play an important part in future work on this issue because its great sensitivity will extend the observa-tions to the magnitude range of +10 to +20, where the individual meteor particles probably have masses of the order of 10^{-6} g and less. Here the numbers of meteors entering the atmosphere are very large indeed—perhaps 10^{19} per day within the sensitivity limit of the telescope. Of course, this estimate is based on an extrapolation from the number distribution of those which are already observed down to +10 magnitude, and one of the most important pieces of information which the new telescope will give is whether this extrapolation is justified. In fact, one of the possible explanations of the zodiacal light would require meteors in this magnitude range to be actually 10 000 times more numerous than in the normal extrapolation. Because of its high definition and gain at wavelengths in the 8 m region, coupled with the steerability of the beam, the telescope is well equipped to study this problem.

Apart from the astronomical problems, the radio echo study of meteors has also given very valuable information about the physics of the high atmosphere* in the 80–120 km region, par-ticularly in the measurement of winds, scale heights, densities and diffusion, etc. The processes by which the meteors evaporate are understood in a very limited sense, and the theories of the scattering of radio waves from the ionized trail have been tested only over a restricted range of wavelengths. In many of these problems the telescope will be needed urgently and will be applied to them from time to time as opportunity permits.

(10) SOME MISCELLANEOUS PROBLEMS IN THE SOLAR SYSTEM

(10.1) Solar Radio Emissions

The study of radio waves emitted by the sun is a major depart-ment of research in radio astronomy,† but it has never seemed likely that the study of these solar radio emissions would occupy any appreciable part of the time of the new telescope. The intensity of the radio emissions associated with sunspots and flares is many thousands of times greater than that from the strongest galactic or extragalactic radio sources, and even in the case of the undisturbed sun the radiation from the corona can be readily studied by sensitive receivers coupled to small radio telescopes. Because of the relative technical ease of this work, the study of these solar radio emissions is widespread throughout the world, and although adequate theories of the sunspot and flare radiation are still lacking, the collection of experimental data is excellent. It is probable that the next major advance in this field will be the production of radio pictures of the sun analogous to the spectroheliograms in optical solar studies. In this kind of development the main instrumental problem is not high gain, but rapid scanning of the solar disc with exceedingly high-definition equipment. The solution probably lies along the

* A short summary of the appropriate problems of the solar parallax and of the *Gegenschein* has been given by LOVELL, A. C. B., and CLEGG, J. A.: 'Radio Astronomy', Chapter 22 (Chapman and Hall, 1952).

† Those interested in a full account of the techniques and results should refer to LOVELL, A. C. B.: 'Meteor Astronomy' (Oxford University Press, 1954).

* A recent account of the present position in these aspects of meteor physics has been given by LOVELL, A. C. B.: 'Handbuch der Physik', Vol. 48 (Springer).

† A very full contemporary account of solar radio astronomy is given by PAWSEY, J. L., and BRACEWELL, R. N.: 'Radio Astronomy' (Oxford University Press, 1955).

lines of the multi-element interferometer systems developed in Sydney rather than in large telescopes whose high gain would be wasted and which would not in any case possess the necessary definition.

On the other hand, there are certain special problems associated with the sun in which the telescope may well be used. For example, the corona is believed to be very extensive at low frequencies (calculations[46] indicate that the effective radius at 38 Mc/s should be 5 photospheric radii). An example of how these extremely tenuous regions may be studied by observing the occultation of a radio star by the sun has already been given by Machin and Smith,[47] who conclude that the emission from a radio star is affected at distances of 10 solar radii. In this type of work the sensitivity and definition of the new telescope would be extremely important when used in conjunction with a smaller radio telescope as an interferometer.

(10.2) The Solar Corpuscular Streams

A solar flare is accompanied by strong radio emissions and often by a simultaneous fade-out of long-distance radio signals, sudden enhancement of atmospherics and a magnetic crochet. These terrestrial effects are probably connected with the intense additional ionization below the E-region caused by the abnormal increase of the appropriate ionizing component in the solar radiation. Although the fade-out is a serious matter for long-distance communication, these instantaneous effects are short lived compared with the disturbances which follow about 30 hours after the flare. In this case the ionosphere is exceedingly disturbed, causing fading and severe distortion of radio signals which depend on ionospheric reflection; there are severe magnetic disturbances and displays of the aurora borealis.

These long-delayed effects are often prolonged for several days and are widely believed to be caused by the impact of particles ejected from the sun at the time of the flare and traversing the intervening sun–earth space at a speed of about 1 600 km/s. This corpuscular stream is believed to consist of fully ionized hydrogen, i.e. of protons and electrons, so that the stream remains neutral. There are, however, many difficulties in connection with the entry of these particles into the atmosphere, since their velocity deduced from the time of travel is quite insufficient to allow them to penetrate to the auroral region (i.e. 120 km above the earth's surface). One prominent contemporary theory was initiated by Chapman and Ferraro[48] and later developed by Martyn.[49] It is that the corpuscular stream is deflected and sets up a ring current round the earth at a distance of several earth radii. It is then suggested that the particles responsible for the auroral and ionospheric disturbances are accelerated in from this ring current along the earth's lines of force. The theory is strongly opposed by Alfvén,[50] who believes that the disturbances in the ionosphere are caused by an electric field in interplanetary space which results from a ring current set up round the earth much closer than in the Chapman and Ferraro theory.

Although the presence of rapidly moving hydrogen atoms has been revealed in the earth's atmosphere by auroral spectra,[51] no one has yet succeeded in providing clear proof of the existence of this stream of particles in the sun–earth space. The existence of the high-sensitivity planetary equipment discussed in Section 8 will provide an opportunity for the telescope to be used in an attempt to detect these streams, or the ring current, by the radio echo technique. This experiment is, of course, a gamble, since there is little guide to the particle density or degree of ionization in the streams. Nevertheless, even a negative result would set a useful limit for comparison with the theoretical estimates of Kahn,[52] and any success would be a key discovery in solar physics and solar-terrestrial relationships.

(10.3) The Aurora Borealis

During the past few years the radio echo technique has bee widely used in the investigation of the aurorae. The typic apparatus for the investigation of meteors in the visual and ne visual range has adequate sensitivity for this purpose, and tl observation of the occurrence of aurorae is no longer limited t the hours of darkness. Many of the phenomena observed a complex, and the radio echo results are not fully understood i relation to the visual appearance of the aurorae, neither is the any satisfactory theory of the mechanism of reflection of tl radio waves. Some of the problems are well illustrated in tl Jodrell Bank aurorae results.[53] For example, the diurn variation of the visual aurorae shows one peak at 22 hour whereas the radio echo data show a minimum at this time ar peaks at 18 and 02 hours. At the time of this minimum, tl drift speeds deduced from the radio echo range movements sho a reversal from E. to W., to W. to E., and reach maxima velocity of about 500 m/s at the time of the maxima in rate occurrence. This change in drift speed is closely correlated wi the variation in the ΔV magnetic component.

No doubt the systematic observations now in progress wi existing apparatus will clarify many of these peculiarities. Tl beam widths, however, are too broad to give radio echo informa tion about particular regions of an auroral display, and it is this respect that the occasional use of the new telescope durir great auroral displays may prove very valuable. The gener future of these radio echo auroral studies is obviously importan not only because of the information about the aurorae, but als in the study of the conditions in the 100–140 km region of tl ionosphere, which is the most frequent height of occurrence.

(10.4) Radio Signals from the Planets

In one respect an important task for the new telescope ha already been fulfilled in the discovery by Burke and Franklin in 1955 that Jupiter occasionally emits radio waves. After th announcement the Sydney and Cambridge groups searched the records for any indication of this phenomenon. The Cambrid records, which were on 81 Mc/s, showed no trace of suc emissions, but the Sydney records on 18·3 Mc/s taken durir 1950–51 showed the effect clearly. The analysis of these recor has been published by Shain,[55] who concludes that the emissic comes from a very small portion of the planet and is probab related to a visually disturbed region in the south temperate be The original discovery was made on a frequency of 22 Mc/s, ar Kraus[56] at Ohio has since studied the signals at 26·6 Mc/s. view of the negative Cambridge results there must be a cut-c somewhere between 26·6 and 81 Mc/s. These experiments rai many questions such as the nature of the spectrum of the rad emission and its origin. Are they violent thunderstorms in tl Jovian atmosphere, or are they in some way connected with tl famous red spot? The opportunities for investigating the Jovia atmosphere and possible ionosphere by such observations a considerable.

More recently Kraus[57] at Ohio has announced the detection radio emissions from Venus, and it seems that a new techniq for planetary investigations is now available. No doubt ma of the problems raised by these planetary radio emissions w be solved by the time there is opportunity to use the new telesco on this type of work, but should occasion demand, the instrume could be quickly applied to these studies over a wide waveleng range.

(11) CONCLUSION

The intention in this lecture has been to review some of tl problems on which the new telescope is likely to be used durir the first few years of its working life. On comparing this accou

vith the original programme drawn up over six years ago, one s impressed not only by the similarity of the programme, but also by the increased importance which an instrument of this ype has assumed during this period of extremely rapid developnent in radio astronomy. Nearly all the original problems of galactic and extragalactic radio emission remain to be solved, and the telescope will enter the field at a moment of extreme nterest. In the radio echo work and other solar system studies he plans for the International Geophysical Year have underlined he significance of the telescope in this sphere. It is confidently hoped that this instrument will make a significant contribution o the International Geophysical Year and thereafter help to naintain the long tradition of observational astronomy in his country.

(12) REFERENCES

(1) Jansky, C. M., Jr.: 94th Meeting of the American Astronomical Society, March, 1956.
(2) Hanbury Brown, R., and Hazard, C.: *Nature*, 1950, **166**, p. 901.
(3) Jansky, K. G.: *Proceedings of the Institute of Radio Engineers*, 1932, **20**, p. 1920.
(4) Jansky, K. G.: *ibid.*, 1933, **21**, p. 1387.
(5) Reber, G.: *ibid.*, 1940, **28**, p. 68.
(6) Reber, G.: *Astrophysical Journal*, 1944, **100**, p. 279.
(7) Mills, B. Y.: *Australian Journal of Physics*, 1955, **8**, p. 368.
(8) Westerhout, G., and Oort, J. H.: *Bulletin of the Astronomical Institutes of the Netherlands*, 1951, **10**, p. 323.
(9) Hanbury Brown, R., and Hazard, C.: *Philosophical Magazine*, 1952, **43**, p. 137.
(10) Baldwin, J. E.: *Observatory*, 1955, **75**, p. 229.
(11) Shklovskii, I. S.: *Astronomical Journal of Moscow*, 1952, **29**, p. 418; 1953, **30**, p. 15.
(12) Shakeshaft, J. R., Ryle, M., Baldwin, J. E., Elsmore, B., and Thomson, J. H.: *Memoirs of the Royal Astronomical Society*, 1955, **67**, p. 97.
(13) Bolton, J. G., Stanley, G. J., and Slee, O. B.: *Australian Journal of Physics*, 1954, **7**, p. 110.
(14) Bolton, J. G., Stanley, G. J., and Slee, O. B.: *Nature*, 1949, **164**, p. 101.
(15) Dombrovsky: *Doklady Akademie Nauk U.S.S.R.*, 1954, **94**, p. 1021.
(16) Shklovskii, I. S.: *ibid.*, 1953, **90**, p. 983.
(17) Oort, J. H., and Walraven, Th.: *Bulletin of the Astronomical Institutes of the Netherlands*, 1956, **12**, p. 285.
(18) Hanbury Brown, R., and Hazard, C.: *Nature*, 1952, **170**, p. 364.
(19) Ryle, M., and Smith, F. G.: *ibid.*, 1948, **162**, p. 462.
(20) Baade, W., and Minkowski, R.: *Astrophysical Journal*, 1954, **119**, p. 206; 1954, **119**, p. 215.
(21) Walsh, D., and Hanbury Brown, R.: *Nature*, 1955, **175**, p. 808.
(22) Hanbury Brown, R., Palmer, H. P., and Thompson, A. R.: *ibid.*, 1954, **173**, p. 945.
(23) Baldwin, J. E., and Dewhirst, D. W.: *ibid.*, 1954, **173**, p. 164.
(24) Hanbury Brown, R., and Hazard, C.: *ibid.*, 1950, **166**, p. 901; *Monthly Notices of the Royal Astronomical Society*, 1951, **111**, p. 357.
(25) Hazard, C.: *Occasional Notes of the Royal Astronomical Society*, 1954, **3**, p. 74.
(26) Baldwin, J. E.: *Nature*, 1954, **174**, p. 320.
(27) Pawsey, J. L., and Bracewell, R. N.: 'Radio Astronomy' (Oxford, 1955.).
(28) Mills, B. Y.: *Australian Journal of Scientific Research*, 1952, A **5**, p. 266.
(29) Bolton, J. G., Stanley, G. J., and Slee, O. B.: *Australian Journal of Physics*, 1954, **7**, p. 110.
(30) Ryle, M., and Scheuer, P. G.: *Proceedings of the Royal Society*, A, 1955, **230**, p. 448.
(31) Reported by Pawsey, J. L., at the Jodrell Bank Symposium, 1955.
(32) Van de Hulst, H. C.: *Nederlandsch Tijdschrift voor Natuurkunde*, 1945, **11**, p. 210.
(33) Ewen, H. I., and Purcell, E. M.: *Nature*, 1951, **168**, p. 356.
(34) Muller, C. A., and Oort, J. H.: *ibid.*, 1951, **168**, p. 357.
(35) Christiansen, W. N., and Hindman, J. V.: *Australian Journal of Scientific Research*, 1952, A **5**, p. 437. (Original announcement by Pawsey, J. L.: *Nature*, 1951, **168**, p. 358.
(36) Van de Hulst, H. C., Muller, C. A., and Oort, J. H.: *Bulletin of the Astronomical Institutes of the Netherlands*, 1954, **12**, p. 117.
(37) Williams, D. R. W., and Davies, R. D.: *Nature*, 1954, **173**, p. 1182.
(38) Heeschen, D. S.: *Astrophysical Journal* (in publication; privately communicated).
(39) Lilley, A. E., and McClain, E. F.: Naval Research Laboratory, Washington, Report No. 4689, December 27, 1955.
(40) Bay, Z.: *Hungarica Acta Physica*, 1946, **1**, p. 1.
(41) De Witt, J. H., and Stodola, E. K.: *Proceedings of the Institute of Radio Engineers*, 1949, **37**, p. 229.
(42) Kerr, F. J., and Shain, C. A.: *ibid.*, 1951, **39**, p. 230.
(43) Browne, I. C., Evans, J. V., Murray, W. A. S., and Hargreaves, J. K.: *Proceedings of the Physical Society B* (in publication).
(44) Murray, W. A. S., and Hargreaves, J. K.: *Nature*, 1954, **173**, p. 944.
(45) Evans, J. V.: *Proceedings of the Physical Society B* (in publication).
(46) Link, F.: *Bulletin of the Central Astronomical Institute of Czechoslovakia*, 1952, **3**, p. 6.
(47) Machin, K. E., and Smith, F. G.: *Nature*, 1952, **170**, p. 319.
(48) Chapman, S., and Ferraro, V. C. A.: *Terrestrial Magnetism and Atmospheric Electricity*, 1931, **36**, pp. 77 and 171; 1932, **37**, p. 147.
(49) Martyn, D. F.: *Nature*, 1951, **167**, p. 92.
(50) Alfvén, H.: 'Cosmical Electrodynamics' (Oxford, 1950); *Tellus*, 1955, **7**, p. 50.
(51) Gartlein, C. W.: *Transactions of the American Geophysical Union*, 1950, **31**, p. 18; Meinel, A. B.: *Astrophysical Journal*, 1951, **113**, p. 50.
(52) Kahn, F. D.: *Monthly Notices of the Royal Astronomical Society*, 1950, **110**, p. 483.
(53) Bullough, K., and Kaiser, T. R.: *Journal of Atmospherics and Terrestrial Physics*, 1954, **5**, p. 189; 1955, **6**, p. 198.
(54) Burke, B. F., and Franklin, K. L.: *Journal of Geophysical Research*, 1955, **60**, p. 213.
(55) Shain, C. A.: *Australian Journal of Physics*, 1956, **9**, p. 61.
(56) Kraus, J. D.: *Sky and Telescope*, 1956, **15**, p. 358.
(57) Kraus, J. D.: *Nature*, 1956, **178**, pp. 33, 103 and 159.

621.395.65

Paper No. 1993]
Mar. 195

TIME SHARING AS A BASIS FOR ELECTRONIC TELEPHONE SWITCHING

A Switched-Highways System

By L. R. F. HARRIS, M.A., Associate Member.

(*The paper was first received* 13*th September, and in revised form* 10*th November*, 1955. *It was published in March*, 1956, *and was read befor the* RADIO AND TELECOMMUNICATION SECTION 9*th May*, 1956.)

SUMMARY

All telephone systems incorporate time-sharing techniques in order to economize in apparatus, but the extent of their application depends upon the operating speeds and adaptability of the components used. The introduction of electronic techniques enables various switching, coding and sampling techniques to be used to achieve a higher degree of time sharing than has previously been possible.

In a system now under development these principles are the main basis of design. Connections are made over common leads using pulse channels each of which may be used for connections to any line. The connections are controlled by register and supervisory apparatus in which information is received, stored, manipulated and transmitted by common apparatus which operates on the connection pulse channels. Common selecting, translating and marking apparatus is used to set up all connections on a one-at-a-time basis, and all apparatus apart from the line-terminal equipment is time-shared by the connections using either sampling or switching techniques. Connection, register and supervisory apparatus is also time-shared by coding.

The capital and maintenance costs of such a system are as yet indeterminate, but it is expected that they will be competitive.

(1) INTRODUCTION

During the past few years the possible application of electronic techniques to automatic telephony has been examined.[1] Although these techniques differ considerably from those of earlier systems, the same underlying principles for achieving an economic design may be applied. It is suggested in the paper that the most important feature which recurs throughout the development of telephone systems is the principle of time sharing of apparatus. Apparatus is time-shared if it is used for different circuits, connections or functions at different times. If the means for associating an item of apparatus with different functions or connections at different times cost less than the items saved, the introduction of time sharing will reduce capital costs. Although this does not imply that a system using more time sharing than another is necessarily cheaper or better, the many examples of time sharing already in use indicate its importance.

(1.1) Time Sharing by Switching

At the beginning of telephony, telephone users had a number of instruments and lines with which each user could call and speak to a number of others. With the realization that no connection was required continuously, switching was introduced which enabled first a single instrument, and later a single line, connected to a central exchange to be used for connections from one subscriber to a number of others on a one-at-a-time basis. Since this introduction of time sharing the main problem of telephone switching engineers has been to provide economic means of interconnection within the exchange, and time sharing has been an important feature of all the solutions proposed. The art of trunking with its associated traffic studies has evolved from the need to time-share trunks, switches and other apparatus among a number of circuits or connections. Common control such as directors and markers provide further examples in whicl a higher degree of time sharing is achieved, since these are hel only during the setting-up of calls.

The possible range of application of switching technique increases with the operating speed of the components used an with their ability to withstand large numbers of operations a high speed. With the introduction of electronic techniques botl the operating and holding times of many items of apparatus ca be much reduced and a higher degree of time sharing achieved For example, a single marker may be used to set up all the con nections in an exchange on a one-at-a-time basis, and expensiv multi-voice-frequency signalling receivers need be connected t a circuit only when there is a signal to be received over it. Botl these examples illustrate time sharing where switching is involved

(1.2) Time Sharing by Sampling

If instead of transmitting the whole of a waveform, sample are taken from which the original waveform or the informatio contained therein may be recovered, a common circuit may b used to transmit the samples relating to a number of connections provided that each connection uses different sampling times The common circuit is used cyclically by the connections and i time-shared by them. The samples for each connection ar transmitted as modulated pulses comprising a pulse train.

The sampling rate required is determined by the rate of chang of the sampled waveform and must be something over twice th maximum frequency to be transmitted. A sampling rate o about 10 kc/s is therefore required for the transmission of speecl signals, and such a rate is possible only with electronic tech niques. The time allotted to each sample must be sufficient t meet transmission and crosstalk requirements, and using presen techniques it cannot be less than about 1 microsec. This set an upper limit of about 100 to the number of speech connection which can be made over a common circuit. In the paper, puls trains and common circuits used for the transmission of speech frequency signals will be referred to as "pulse channels" an "highways," respectively, and it is convenient to refer to connec tions using coincident pulse trains as using the same puls channel.

Such sampling or time-division-multiplex techniques may als be used to transmit other signals, and systems have been describe in which the detection of dial pulses and other line signals i achieved by sampling the line signalling condition.[3,4] Here a sampling rate of 60 c/s is adequate since the relevant wavefor changes occur more slowly.

Since much of the information used to set up and contro connections in an exchange is presented transitorily, memor devices are essential and the storage of information is an im portant aspect of any telephone system. Among the electroni storage devices are several in which information is stored as th presence or absence of pulse trains. The apparatus is time shared by the bits of information stored, and such storage device

Mr. Harris is in the Post Office Engineering Department, Research Branch.

nave capacities equal to several hundreds or even thousands of elays. For example, magnetic-drum tracks, magnetostriction and mercury delay lines may store 3 000, 1 000 and 300 pulse rains having a pulse repetition frequency of 60 c/s, 1 kc/s and 0 kc/s, respectively. These figures are representative, and the apacities and operating speeds will no doubt increase as new and improved techniques become available. In order to use hese stores efficiently, the rate at which the information is presented should be related to the sampling rates required for the nformation which they store or control. In order to control the ransmission of speech samples, delay lines of 100 microsec using oulse trains of 10 kc/s repetition frequency are appropriate, whereas for the storage of dialled information magnetic drums may prove more applicable.

In electro-mechanical systems the progress of a call is controlled by opening and closing contacts which route speech and other signals along paths determined by presented and stored nformation. In time-division-multiplex systems a call may be controlled by opening or closing electronic gate circuits, which oute speech and signal samples along paths determined by the nformation applied to the gates as the presence or absence of oulses. The gate circuits are electronic equivalents of contacts out are time-shared over the connections by sampling.

(1.3) Time Sharing by Coding

Much of the apparatus used in electro-mechanical systems can be in either of two states. Thus relays can be operated or not, contacts can be open or closed, etc., and using such binary elements any sequence of operations can be controlled. It is common practice to use a group of relays to perform a number of different functions depending upon the combinations of relays operated. Thus by suitably coding the information relating to a connection, a relay or contact may be used for different functions at different times. Current techniques show many examples where time sharing is achieved in this way, and various theoretical aids such as Boolean algebra and communication theory have been used in the development of efficient coding systems.

These same theoretical aids may be applied when electronic equipment is being designed. Again, any sequence of operations may be controlled using such binary elements as the presence or absence of a pulse train in a store and the open and closed conditions of a gate circuit, and a high degree of time sharing by coding can be achieved even when the apparatus is already time-shared by sampling.

(1.4) Time Sharing in Electronic Systems

While the introduction of electronic techniques greatly extends the range of application of time-sharing principles, they introduce no change in the approach to the basic problem of designing a telephone system using a minimum of apparatus for a given service. The newer techniques merely make possible the wider application of the solution to the problem—the solution which has been with the art since the beginning.

During the past few years various time-sharing techniques have been under development which can be applied in varying degree to many switching systems. One such technique involves switching between pulse channels on different highways, and systems embodying this are called "switched-highways systems." A particular arrangement has been selected for detailed study because time sharing can be applied to a greater extent in this system than in any other envisaged at the present time. This does not necessarily make it the cheapest to build or maintain, but it serves to illustrate some time-sharing techniques which are likely to find wider application in the future. In the paper this system will be referred to as the switched-highways system.

(2) CONNECTION PATHS

(2.1) Switched-Highways Trunking

In the switched-highways system all the external lines—subscribers' lines, junctions, manual-board lines, etc.—are arbitrarily arranged in groups which carry about the same total traffic and have about the same number of lines. Each line in a group is connected by line gates to a pair of highways used one for each direction of transmission, which are common to all the lines in the group, and the highways of each group are connected by other gates to the highways of every other group. In Fig. 1 the highways of the five groups GP1–5 are depicted by single lines H1–5 and the highways H1 of GP1 are connected by line gates LG to each of M lines of which one, L1, is shown. Highways H1 are also connected to highways H2–5 by gates G1/2, G1/3, G4/1 and G5/1, respectively.

The pulse channels on the highways of each group are made available to all the lines connected to the group, and a line may

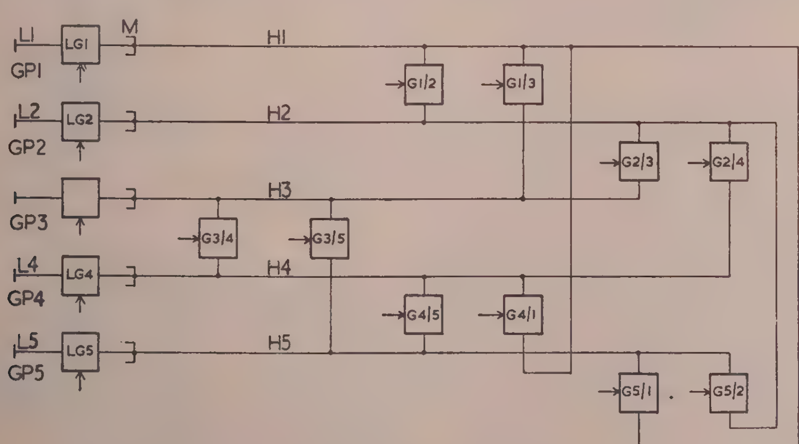


Fig. 1.—Switched-highways system: basic trunking.

GP—Group.
L—External line.
LG—Line gates.
H—Group transmit and receive highways.
G—Inter-group gates.
M—Number of external lines in GP1.
N—Number of pulse-channels.

be connected to any of the pulse channels by applying appropriate pulses to its line gates. The same pulse channels are used on all the highways, and two lines in different groups are connected together by connecting them to the same pulse channel and by applying pulses coincident with this channel to the gates interconnecting the highways of their groups. Thus L1 in GP1 can be connected to L2 in GP2 by applying pulses appropriate to one pulse channel, PC1 say, to LG1, LG2 and G1/2. While this pulse channel PC1 is in use for this connection it cannot be used for other connections to lines in GP1 or GP2. However, it can still be used for connections between lines in the remaining groups. Thus L4 in GP4 may be connected to L5 in GP5 by pulsing LG4, LG5 and G4/5 with PC1 pulses. The two connections will not mutually interfere because no pulses appropriate to PC1 are applied to the gates connecting H1 or H2 with H4 or H5. With this arrangement any of N pulse channels can be used to connect any pair of lines in different groups, provided that the pulse channel is not already in use on either of the highways of their groups.

The connection of two lines in the same group presents problems which are discussed later.

This basic trunking is a time-division-multiplex equivalent of the electro-mechanical system shown in Fig. 2, in which the

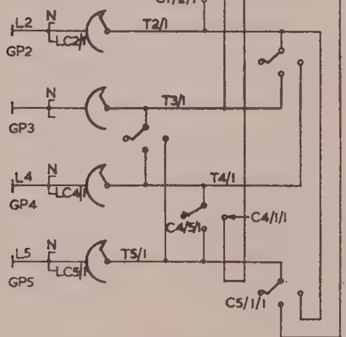


Fig. 2.—Electro-mechanical equivalent of switched-highways-system basic trunking.

GP—Group.
L—External line.
LC—Line contacts.
T—Trunk.
C—Inter-trunk contacts.
M—Number of external lines in GP1.
N—Number of trunks.

highways of each group are replaced by N trunks which are fully available to the lines in the group and in which corresponding trunks in the groups are interconnected by contacts. Thus T1/1 of GP1 is connected by contacts LC to the M circuits in the group, one of which, L1, is shown. T1/1 is also connected to T2/1, T3/1, T4/1 and T5/1 by contacts C1/2/1, C1/3/1, C4/1/1 and C5/1/1, respectively. L1 and L2 may be connected by closing LC1/1, LC2/1 and C1/2/1, and at the same time L4 and L5 may be connected by closing LC4/1, LC5/1 and C4/5/1. In the arrangement of Fig. 2, the trunks are time-shared by making them available to all the lines. In Fig. 1 the apparatus is further time-shared by sampling, so that each pulse channel is equivalent to an assemblage of interconnected trunks. This reduces the number of switching elements in the ratio $N:1$, but this is perhaps an over-simplification of the comparison.

(2.2) Basic Switches

Fig. 1 illustrates two types of switches. The line gates of a group connect lines separated in space to pulse channels separated in time. They therefore make up an $M:N$ switch in which N pulse channels carry the both-way traffic of M lines. The traffic-carrying capacity of the switched-highways system is discussed in Appendix 11. With, say, 80 pulse channels available the traffic per group may be about 52 erlangs. The number of lines, M, may therefore be considerable, varying from about 150 to 1 000, depending upon the traffic per line. The second type of switch occurs on the highways and connects pulse channels separated in space. Ten such switches are shown in Fig. 1 and each is equivalent to N separate $1:1$ switches. There is no theoretical limit to the number of groups, and the practical limit appears to be very large. The number of $N(1:1)$ switches required if the highways are interconnected directly is equal to the number of combinations of G taken two at a time, i.e. ${}_GC_2$, where G is the number of groups. If G is very large, some economy may be achieved by interconnecting the highways in stages using multiplexed equivalents of well-known electro-mechanical arrangements. However, the saving is comparatively small unless G exceeds 50, and this number would be adequate for most 10 000 line exchanges.

(2.3) Other Switched-Highways Trunking Arrangements

These two types of switches may be incorporated in a variety of other systems which cannot be described here in detail. For example, each group of lines may be provided with two sets of highways, one used for incoming and the other for outgoing traffic. In this arrangement, line gates connect each both-way external line to each set of group highways and other gates interconnect the incoming highways of every group with the outgoing highways of every group. Two external lines are then connected using the same pulse channel on the incoming highways of the incoming line's group and the outgoing highways of the outgoing line's group. The highways would carry unidirectional traffic and each group could contain twice the number of lines, but the saving in the number of groups is unlikely to justify the provision of two sets of line gates for each both-way line, although the arrangement might find application in trunk exchanges if most of the lines carried unidirectional traffic.

Other variations on the switched-highways trunking are illustrated in Fig. 3, which shows one of a number, Q_1, of groups

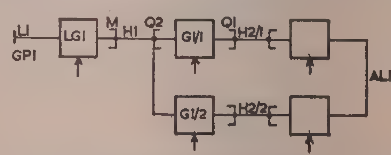


Fig. 3.—Alternative trunking using a.f. links.

L—External line.
LG—Line gates.
G—Inter-highway gates.
H1—Group transmit and receive highways.
H2—Secondary transmit and receive highways.
AL—A.F. links.
M—Number of external lines in GP1.
Q_1—Number of groups of external lines.
Q_2—Number of secondary highways.

of lines provided with highways H1. The highways of each group of lines are interconnected by gates to each of a number, Q_2, of secondary highways of which two, H2/1 and H2/2, are shown. These secondary highways are interconnected by a.f. links to which they are connected by gates. Thus AL1 would be one of the a.f. links used to connect H2/1 and H2/2. Each connection is now made using an a.f. link, the incoming line being connected to one end of the link using one pulse channel, and the outgoing line being connected to the other end of the link using another pulse channel.

So far it has been assumed that any of the pulse channels may be transmitted through any of the gates by applying suitable

ulses from memory apparatus used to remember the connections hrough the exchange. With the arrangement of Fig. 3, addiional gating stages have been introduced on each connection nd an adequate number of possible paths may be provided etween any pair of lines without making all the pulse channels vailable for all the possible routes through the exchange. It is herefore possible to economize in memory apparatus by fixing he pulse channels used at one of the three gating stages. For xample, it could be arranged that the use of a particular a.f. link vould require the use of particular channels. Alternatively, it vould be possible to fix the pulse channels used by the external ines. The line gates would then be used merely to translate the ignals on the space-separated lines into signals separated in ime. From a switching point of view these gates would be vasted, but nevertheless a system might be developed on this asis which would be comparable in cost to the switched-highvays system itself. Alternatively, it would be possible to fix the ulse channels transmitted through each inter-highway switch, hus limiting the number of connections which could exist etween any group of lines and a secondary highway. Each of hese alternatives might find application in certain circumstances, ut attention has been concentrated on the switched-highways ystem itself because—since no a.f. links are involved—a higher egree of time sharing is achieved.

The allocation of particular pulse channels for particular interroup connections is directly applicable to the basic switchedighways trunking if the number of groups is small. In Fig. 1, or example, using 100 pulse channels PC1–100, PC1–20 could e used for connections through G1/2 and G3/4; PC21–40 hrough G1/3 and G4/5; PC41–60 through G4/1 and G5/2; C61–80 through G5/1 and G2/4; and PC81–100 through G2/3 nd G3/5. Only 20 connections could then exist between any air of groups, and it is unlikely that the saving in memory pparatus would justify the loss in availability even in this xample; with large numbers of groups the technique would be ven less worth incorporating.

(2.4) Connection-Path Apparatus

The line gates and highway apparatus used in the switchedighways system are illustrated in Fig. 4, which shows a typical rrangement of the elements required for a 2-wire subscribers' ne. The 2-wire line L is converted to 4-wire by a hybrid ansformer, and is connected to the transmit highway H1/1 y a transmit gate TG and modulator TM. The line-signalling condition is extended over an auxiliary path TS to control the transmission of pulses from TP to H1/1 through TG via the modulator TM in which they are modulated by the signals on the transmit channel. A modulated pulse channel on H2/1 is gated in the receive gate RG by applying coincident pulses on RP, and the modulating signal is recovered in the unit DM, which may include an amplifier and low-pass filter. A receive signalling condition on RS may be derived in RR from the receipt of pulses in DM. By applying suitable pulses on TP and RP the line may be connected to any pulse channel.

The transmit and receive highways H1/1 and H2/1 are connected by gates such as G1/2 and G2/1 to the receive and transmit highways, respectively, of all the remaining groups. These inter-group gates are conveniently situated at a central point, and unless the number of groups is very small the highways will be sufficiently long to require the use of coaxial cable. Cable-drive and terminating units will be required as illustrated in Fig. 4, and in order to overcome cable delays without reducing the number of channels, it becomes necessary to delay the pulses applied to the line receive gates with respect to those applied to the line transmit gates. If t is the cable delay, then, in order to make connections between lines in any pair of groups, the pulse channels on the highways at the highway-switch end must occur coincidently on all highways. Therefore if a channel at this point coincides with pulses P_x, pulses P_{x-t} and P_{x+t}, occurring t before and after P_x, must be applied to the line transmit and receive gates.

The introduction of cable into the transmission system may reduce the number of pulse channels which can be used in order to meet the crosstalk and transmission equivalent requirements. It is estimated that adequate performance can be obtained using 80 channels with a 10 kc/s sampling rate and a channel time of 1·2 microsec.

(2.5) Within-Group Connections

So far, no method of making connections between lines in the same group has been described. If one pulse channel only is used for a within-group connection the pulse sample from a line's transmit modulator is gated by the same line's receive gate, which results in unacceptable side-tone. It is therefore not practicable to use only one pulse channel for within-group connections, and two must be used together with means for interconnecting them. It would be possible to use a.f. links in the same way as those shown in Fig. 3, but solely for

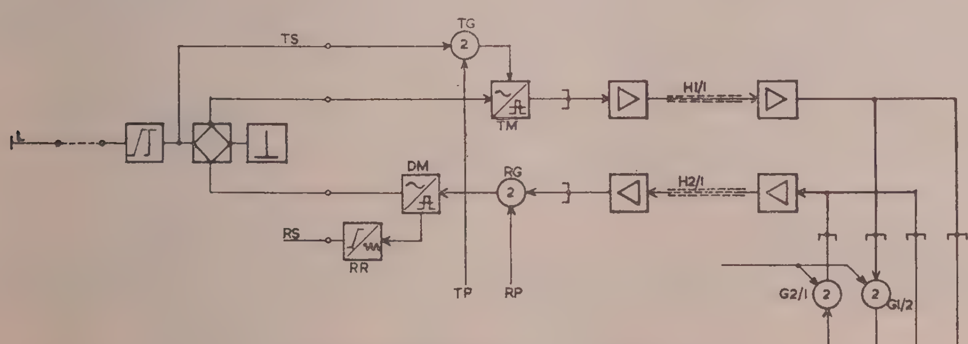


Fig. 4.—Connection path apparatus.

L—External line.
TS—Transmit signalling path.
TG—Transmit gate.
TM—Transmit modulator.
H1/1—Transmit highway.
H2/1—Receive highway.
RG—Receive gate.
DM—Demodulator and amplifier.
RR—Rectifier.
RS—Receive signalling path.
TP—Transmit-gating-pulse lead.
RP—Receive-gating-pulse lead.
G—Inter-group gates.

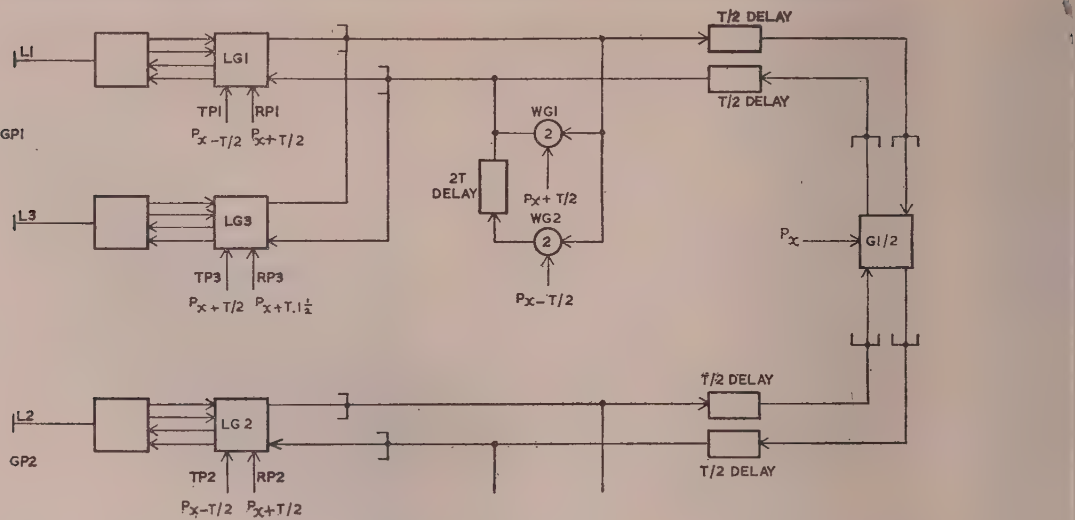


Fig. 5.—Within-group connections.

GP—Group.
L—Line.
LG—Line gates.
TP—Transmit-gating-pulse lead.
RP—Receive-gating-pulse lead.
G—Inter-group gates.

WG—Within-group gates.
T—Pulse-channel duration.
$P_{x-T/2}$, P_x, $P_{x+T/2}$—Phases of pulse trains separated by $T/2$ and appropriate to pulse channel PCx.
$P_{x+T/2}$, P_{x+T} $P_{x+T\cdot1\frac{1}{2}}$—Phases of pulse trains separated by $T/2$ and appropriate to pulse channel PC(x + 1).

within-group connections. As shown in Appendix 11.3, one secondary highway connected to all the a.f. link terminations would suffice for any size of exchange.

An alternative technique[5] in which no a.f. links are required is shown in Fig. 5. If the delay of each cable is made equal to half the channel time T, a delay of T is required between the pulses applied to the line transmit and receive gates. Connections between lines in different groups can be made as before using one pulse channel, but with a total delay T over the highways. Thus in Fig. 5, L1 and L2 may be connected using PCx by applying P_x to G1/2, $P_{x-T/2}$ on TP1 and TP2 and $P_{x+T/2}$ on RP1 and RP2. If two lines in the same group, e.g. L1 and L3, use adjacent pulse channels PCx and PC(x + 1) respectively, they may be connected by causing the output of the transmit modulator of L1, occurring at $P_{x-T/2}$, to be delayed by $2T$ and presented to the receive gate of L3 at $P_{x+T.1\frac{1}{2}}$, and by causing the pulse transmitted at $P_{x+T/2}$ from L3 to be received after negligible delay by L1. This is achieved using gates WG1 and WG2, which connect the line ends of the highways together directly and via a delay of $2T$ when pulsed with $P_{x+T/2}$ and $P_{x-T/2}$, respectively. Any pair of adjacent channels may therefore be used for within-group calls, and a high degree of time sharing is again achieved.

(2.6) Register Connections

As automatic telephone networks have become more complex there has been an increase in the use of registers to control the setting up of connections, and it is in the use of registers anc other complicated control apparatus that the principle of time sharing has been carried furthest. With high-speed electronic equipment this application of the principle can be extended, anc register-marker exchanges using a single marker to set up al calls on a one-at-a-time basis become possible even with the largest installations.

The switched-highways exchange is a register-marker system in which a single marker is used and in which the register apparatus is time-shared not only by the external lines using switching techniques but is also time-shared by sampling. The highways of each group are connected by gates to register highways which are used to transmit called-line identification and other signals between the register apparatus and the lines. The number o register highways required will depend upon the total registe traffic, since it is impossible to carry more connections on a highway than there are pulse channels. By way of illustration it is convenient to consider an exchange in which the number o registers is equal to the number of pulse channels, and two sets o highways are provided, one for incoming traffic and one for outgoing. Such an arrangement appears in Fig. 6, which shows the highways of one group interconnected with the incoming registe highways IRH and the outgoing register highways ORH by gates G1/R1 and G1/R2, respectively.

When a line calls it is connected to the register apparatus using a pulse channel which is free on its own group and on the incoming register highways. This pulse channel is used to senc

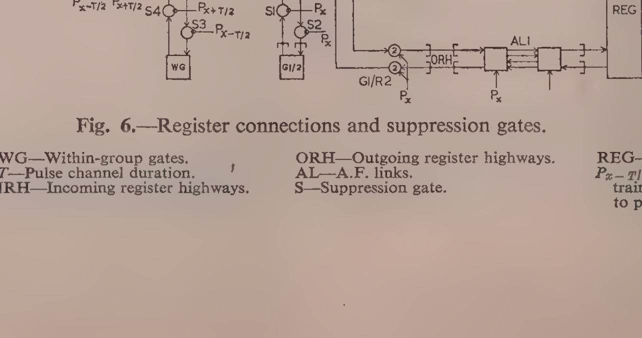


Fig. 6.—Register connections and suppression gates.

L—Line.
LG—Line gates.
G—Inter-highway gates.

WG—Within-group gates.
T—Pulse channel duration.
IRH—Incoming register highways.

ORH—Outgoing register highways.
AL—A.F. links.
S—Suppression gate.

REG—Register apparatus.
$P_{x-T/2}$, P_x, $P_{x+T/2}$—Phases of pulse trains separated by $T/2$ and appropriate to pulse channel PCx.

and receive signals over the incoming line and is also used in the register apparatus to store and manipulate the designation and other information relating to the call.

When sufficient designation information has been received the marker is brought in to attempt to set up the outgoing connection. On some calls a connection between the outgoing line and the register apparatus must be established in order to transmit and receive signals which will further the progress of the call. When such an outgoing register connection is required, mutual interference between the incoming and outgoing register connections must be avoided, and it is therefore necessary to prevent the transmission of signals between the incoming and outgoing lines until the register releases. It is also necessary, however, to check that this connection between the lines can be made before or at the same time as the outgoing register connection is made. Both these requirements are met if the two connections to the called line are set up at the same time and if they use the same pulse channel. The transmission of signals between the incoming and outgoing lines is then prevented if pulses of this channel are derived from the outgoing register connection and applied to suppression gates inserted in the group highways at points which do not interfere with the register connections, i.e. at points marked S in Fig. 6, on the inputs and outputs of the within-group gates WG and inter-group-highway gates such as G1/2. Thus if PCx is used for an outgoing register connection to L1 in Fig. 6, P_x is applied to S1 and S2, and $P_{x-T/2}$ and $P_{x+T/2}$ are applied to S3 and S4, respectively, until the register releases. If PCx is used for the connection between L1 and a line in a different group, transmission between lines is prevented at S1 and S2. If PCx is used on a within-group connection, the connection is inhibited at S3 and S4. When this technique is used, it is necessary to set up a connection between lines in different groups using a channel which is free in both groups and is also free on the outgoing register highways. This restriction on the choice of pulses will reduce the traffic-carrying capacity of the highways, but if the traffic carried on the outgoing register highways is small, or if the outgoing register traffic can be shared over a number of sets of highways thus giving a choice of highways on which to find a suitable free channel, the effect is not serious.

In order to send and receive information, which is stored in the register apparatus using the pulse channel used for the incoming line-to-register connection, over the outgoing line which is using a second channel, a.f. links such as AL1 are inserted between the register apparatus and the outgoing register highways. The number of a.f. links required is determined solely by the outgoing-register traffic.

On many calls no outgoing register connection is required and the marker merely operates to effect the connection between lines. In general, on such calls the register will be released as soon as the marker has set up the connection.

(3) CONNECTION MEMORY APPARATUS

In order to set up and maintain connections in a telephone exchange the basic functions of selection and memory are required. In electro-mechanical systems these two functions are so closely related that they are often performed by the same mechanism, whereas in the switched-highways system they are separated and the lines and channels to be used in a connection are selected by common apparatus and indicated to slave storage devices.

In this system the stored information must be presented to the gates in the form of pulses. Each gate could be connected to a pulse store which generates just those pulses required by the gate. Alternatively and more economically, the stores may be time-shared by coding, so that the effectiveness of a pulse in any particular gate depends upon its coincident generation by a particular combination of pulse stores. Any pulse train may then be made effective in a gate by storing it in the gate's combination of stores. With P pulse stores there are $2^P - 1$ ways of generating a pulse by at least one of them, and therefore in principle $2^P - 1$ different gates may be supplied by them. However, if only pulses generated by a particular combination of stores are to be effective in a gate, additional gating apparatus must be provided. In order to simplify these gates it may be convenient to restrict the combinations used in practice to something less than the theoretical maximum. For example, by using the combinations of P stores taken a fixed number, R, at a time, this gating apparatus may be the same for all the ${}_PC_R$ gates which may be served. Fig. 7 shows a simple arrangement in which four pulse stores serve six gates which are each connected to two of the stores. Similarly 12 stores may supply 220 gates, and 20 stores may supply 1 140 gates if combinations taken three at a time are used. In practice, these coincidence gates may be incorporated in the speech gates which they serve.

Such a group of stores may be used to supply a group of gates provided that no pulse is required to be effective in more than one gate. In the switched-highways system there are a number

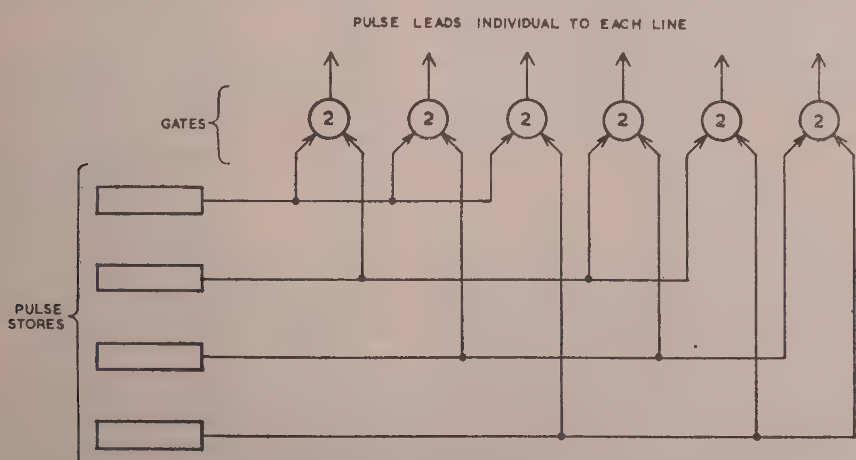


Fig. 7.—Pulse distribution.

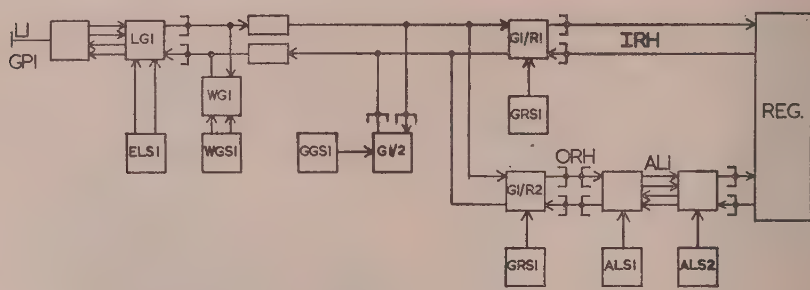


Fig. 8.—Connection stores.

GP—Group.
L—Line.
LG—Line gates.
IRH—Incoming-register highways.
ORH—Outgoing-register highways.
AL—A.F. links.
REG—Register apparatus.
G—Inter-highway gates.
WG—Within-group gates.
ELS1—External-line stores of GP1.
WGS1—Within-group store of GP1.
GGS1—Group-to-group connection store of GP1.
GRS1—Group-to-incoming-register connection store.
GRS2—Group-to-outgoing-register connection store.
ALS1—A.F.-link-to-outgoing-register connection store.
ALS2—A.F.-link-to-register connection store.

of such groups of gates which are shown in Fig. 8 in which the stores ELS1 serve the line gates of the group shown, GRS1 and GRS2 serve the gates which connect the group highways to the incoming and outgoing register highways respectively, GGS1 serves the gates which connect the highways of the group shown to the highways of some other groups, and ALS1 and ALS2 serve the gates connecting the a.f. links to the outgoing register highways and to the register apparatus. It would be possible to use a combination of the GGS1 stores for the within-group gates shown, but it is more convenient to consider these as being served by an individual store, WGS1.

Each store must now supply all the pulses which are required to be effective in all the gates to which it is connected. In general, therefore, it will generate a number of pulses in each cycle, and storage devices of the pulse-circulating type are particularly applicable for this purpose. In order to supply the line speech gates for each direction of transmission using the same stores, different phases of pulses are required. This is conveniently carried out in the way shown in Fig. 9, which shows a delay-line circulating system consisting of a delay $(N - 1)T$ in series with two short delays of $T/2$. The three output points will produce pulses which are separated by $T/2$, and phases A and C may be used to distribute pulses to the line transmit and receive gates, respectively, and phase B of other stores may feed the highway gates. A pulse train is stored by applying a pulse of the train on lead O, which will cause the generation of the pulse train until its circulation is inhibited by applying another coincident pulse on I. Some variation on the timing of these operating and inhibiting pulses and in the length of the delays is tolerable if suitable retiming elements are incorporated. The $T/2$ delays are conveniently provided using electric delay lines, but for a pulse repetition time of 100 microsec, the longer delay of about 99 microsec must be provided by a magnetostriction or mercury-delay line.

(4) SETTING UP CONNECTIONS

(4.1) Selection

The operation of selecting one from a number of equally suitable circuits is a basic requirement of any telephone-exchange system. In the switched-highways system, two setting-up processes are required on each completed call. An incoming line is first connected to the register apparatus, and this process involves two selections. Since more than one line may call at any instant, a selection of the calling line must be made, and in addition, one of the pulse channels which is free in both the calling line's group and on the incoming register highways must be selected. The selected pulse channel must then be stored in the combination of stores associated with the selected calling line and in the

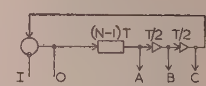


Fig. 9.—Pulse store.

O—Operate lead.
I—Inhibit lead.
T—Pulse channel duration.
A, B and C-outputs separated by $T/2$.
$(N - 1)T$—Delay provided by magnetostriction or mercury delay line.

combination of stores used to control the gates connecting tl selected line's group to the incoming register highways. similar if somewhat more complicated process is required whe the connection to the outgoing line is made.

If each of a number of circuits is characterized by a pulse tra and if the pulse trains of those circuits from which a selection to be made are generated onto a common lead, the problem selecting a circuit becomes the same as that of selecting a pul channel, and the same techniques can be used for both types selection. The problem is solved by any device which, whe presented with a sequence of pulses, operates to one of the and transmits it to an output lead and prevents the subseque transmission of any other pulses. The single pulse appearir on the output lead is then characteristic of the selected channe line or other circuit. In general, some temporary storage of tl selection made will be required before the connection memo apparatus can be operated, and the form of this tempora storage depends upon the type of indication required.

Typical arrangements are shown in Fig. 10, where the pul

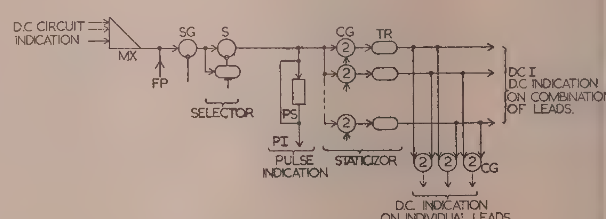


Fig. 10.—Typical selector.

MX—Signalling multiplex.
FP—Pulse input.
SG and S—Suppression gates.
T and TR—Triggers.
PS—Pulse store.
PI—Pulse indication.
CG—Coincidence gates.

trains applied to the selector are either generated by a multipl MX used to convert d.c. indications on leads individual to circui into pulse trains on the common output lead or are appli directly over the lead FP. Unsuitable pulse trains are delete by applying coincident pulses to the suppression gate SG, who output is connected to the selector, which consists of a trigger

ınd a suppression gate S. The first pulse transmitted through S operates T, which then closes S to any further pulses. This pulse operates the pulse store PS which generates the selected pulse rain on the lead PI. If a d.c. indication of a selected circuit is equired, the transmitted pulse is applied to a staticizor in which coincidence gates are used to operate a combination of triggers TR, which is individual to the selected pulse train and therefore to he selected circuit. This staticizor indicates the selected circuit on a combination of d.c. indicating leads DCI which could correpond, for example, to the combination of pulse stores to be used by a selected line. The leads DCI may also be connected to coincidence gates which give an indication of the selected circuit on a lead individual to it. When the connection has been set up, he selector, together with its temporary stores, may be reset and nade available for other selections.

In selecting lines it is convenient to use two stages of selection. One selector selects a group of lines, at least one of which is uitable for selection, and a second selector then selects a line within the selected group. Two selectors may operate to select a line in about 2 millisec. With such high operating speeds it is clear that the same selectors may be used for all connections to be et up through the exchange, and, in fact, it would be possible for a single selector to perform every selection process required.

(4.2) Incoming Line-to-Register Connections

In order to illustrate these techniques the setting up of a connection between a calling line and a register will be described with reference to Fig. 11, which shows one group, GP1. Each group of external lines is provided with three signalling multiplexes in which each line is characterized by a pulse train. Multiplex CGM1 is used to indicate the forward signalling or calling condition of the external lines in GP1, multiplex HM1 indicates their backward signalling condition, and multiplex CSM1 indicates their class of service, as will be described later. The same signalling pulse trains are used in all multiplexes in all groups.

The pulse channels already in use on the group and incoming register highways, IRH, are derived from the stores ELS1 and GRS1, and are indicated to unit U2 over GBP1 and RBP1, respectively. U2 indicates to unit U1/1 over the lead FCP1 those channels which are free on both of those highways. When a line calls, its pulse train is generated by CGM1, which is connected via suppression gate SG1 and lead CGP1 to U1/1. If pulses appear on both FCP1 and CGP1, a d.c. signal is derived in U1/1 indicating that the group includes at least one connectable calling line. Common group selector GPS selects one such indicated group and d.c.-indicates it on a combination of leads corresponding to the combination of pulse stores in GRS1 used to connect its highways to IRH. The selected group is also indicated on a lead individual to the group—lead GL1 for GP1. If GP1 is the selected group, the signal on GL1 opens gate CG1/1, transmitting the free channel pulses on FCP1 to the common channel pulse selector CPS. The signal on GL1 also opens gate CG2/1 transmitting the pulse trains of the calling lines in the selected group to the common line selector CTS. The selected line is d.c.-indicated on a combination of leads corresponding to

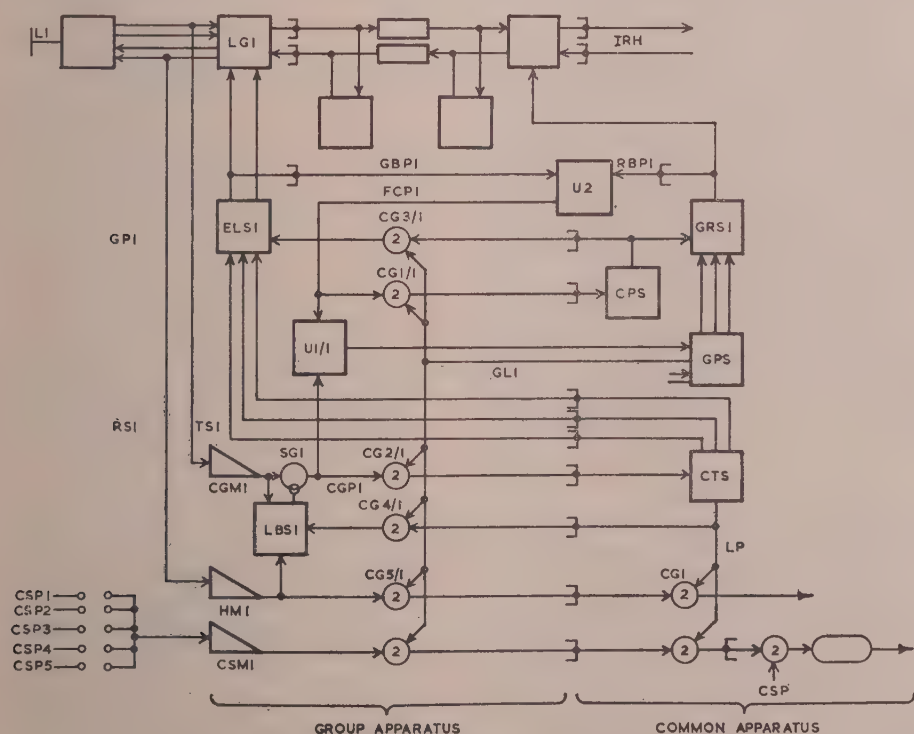


Fig. 11.—Selection for calling-line-to-register connections.

GP—Group.
L—Line.
LG—Line gates.
IRH—Incoming-register highways.
ELS1—External-line stores for GP1.
GRS1—Group-to-incoming-register connection store.
TS1—Forward signalling path of L1.
RS1—Backward signalling path of L1.
CGM1—Forward signalling multiplex of GP1.
HM1—Backward signalling multiplex of GP1.
CSM1—Class of service multiplex of GP1.
CSP—Class of service pulse trains.
LBS1—Line busy store of GP1.
GPB1—Busy channels in GP1.
RBP1—Busy channels on IRH.
FCP1—Common free channels for incoming register connections from GP1.
U—Comparison unit.
CGP—Calling line pulses.
SG—Suppression gate.
CG—Coincidence gate.
GPS—Group selector.
CPS—Channel pulse selector.
CTS—External-line selector.
GL1—Selected group indication for GP1.
LP—Pulse indication of selected line.

the combination of stores used to connect the line to its group highways. The selected channel pulse operates the indicated group-to-register stores GRS1, and is also transmitted through CG3/1, which is opened by the signal on GL1 to operate the indicated external-line stores ELS1 in the selected group. The incoming line-to-register connection is thus established.

The selected line is also pulse-indicated on LP, which is connected via CG4/1 to the "line busy" store LBS1, in which are stored the signalling pulses of all the busy lines in GP1. These pulses are applied to SG1 in order to prevent the reselection of a busy line. The storage in LBS1 is maintained using the outputs of CGM1 and HM1 which indicate the forward and backward signalling conditions of the line, respectively. When the forward signalling condition is extended over the transmit highway to the register using the selected channel, a backward signalling condition is reverted over the group receive highway, and the resultant d.c. condition on RS1, derived from the receipt of pulses in the line receive gate, causes HM1 to generate the line's signalling pulse. The output of HM1 gated through CG5/1 by the signal on GL1 may be compared with the selected-line pulse on LP in gate CG1. The resulting output indicates that the connection between line and register has been satisfactorily set up and may therefore be used to release the selectors, which then become free to set up other connections.

(4.3) Connections to Outgoing Lines

When an outgoing connection is to be set up, the marker operates to indicate the external line or lines which are suitable for the connection, and a selection of one of these is made using techniques similar to those described in Section 4.2. The marked lines in a group are indicated using a marking multiplex, which in practice may be the backward signalling multiplex shown in Fig. 11. The incoming line's group and combination of pulse stores must be identified in order to complete the connections. This is effected using the pulse channel used for the incoming register connection, since this channel is used only on an incoming register connection to one group and is stored only in the incoming line's combination of stores in that group. Again a pulse channel—or pair of channels for within-group calls—must be selected which is free on the appropriate highways. These highways will include the outgoing register highways if the outgoing register connection is required, when a free a.f. link must also be selected. Space does not permit a more detailed discussion of the setting-up process required at this stage of the call, and it must suffice to say that the same group, line and pulse-channel selectors may be used as already described for the incoming connections.

(5) CONNECTION CONTROL

So far, the description of the switched-highways system has been restricted to the provision of speech paths and the means for setting up and remembering the connections. It has been shown that these problems may be solved using a high degree of time sharing, with resultant economy in the amount of apparatus required. In addition, however, many control operations must be performed which involve the reception and transmission of information over the established connections. The register functions will include the reception, storage, recoding and retransmission of information over the external lines, and when the register has released, supervisory apparatus must control the connection between the incoming and outgoing lines.

(5.1) Control Techniques

Many of these functions may be performed by apparatus which is time shared by sampling and operates directly on the pulse channels used for the connections over the highways. F example, a tone may be sent to a line by modulating the puls applied to its line gate. The same tone may be sent to a numb of lines by inserting the pulse channels which they use into pulse store whose output pulses are modulated in a commo modulator and transmitted over the highway to the lines. Su modulators are used in the switched-highways-register apparat for the transmission of dial-tone and inter-register signa Digital information using a two-out-of-five voice frequency co is sent out over the outgoing register highways using five store each of which controls one of the signalling tones. Similar each group of lines is provided with a number of modulators f the sending of n.u., busy, ring and other supervisory and juncti signalling tones.

D.C. information, e.g. dial pulses, may be transmitted to t line by interrupting the pulses applied to the line's receive ga in order to cause corresponding interruptions on the line backward signalling path.

For both the transmission and reception of coded informati various timing operations are required. These can be carri out by comparing the duration of the signal with the tir taken to count a given number of regularly spaced timi pulses and using apparatus which is common to all the pul channels. This is illustrated in Fig. 12, where the function the B relay is taken as an example.

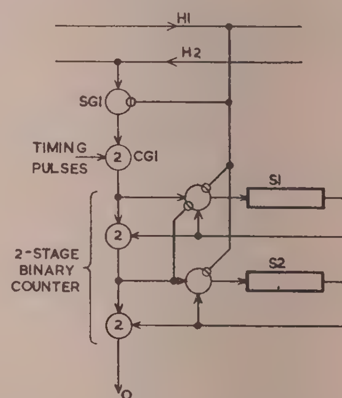


Fig. 12.—Timing operations: B relay function.

H1—Line signalling condition.
H2—Backward signalling condition.
S—Store.
SG—Suppression gate.
CG—Coincidence gate.
O—Output when four timing pulses h been counted.

H1 and H2 are highways indicating the forward and backwa line-signalling conditions, respectively. The pulses on H2 a applied to the suppression gate SG1, in which they are inhibit by coincident pulses on H1. When the forward line-signalli condition on H1 is removed, the pulses on H2 are transmitt through SG1 to the gate CG1, where they are gated by timi pulses which coincide with each channel pulse once every cy of, say, 60 millisec. The pulses of each channel so gated a counted in the binary counter consisting of the stores S1 and S The first gated pulse is stored in S1, the second deletes it fro S1 and stores it in S2, and so on. The fourth pulse transmitt to the counter will give an indication on the output O that t forward signalling condition has been removed for at least thr repetition times of the coincidence between the channel a timing pulses, i.e. at least 180 millisec. This pulse may be used release the connection. If pulses reappear on H1 during such count they reset the counter to zero. These timing techniques a

milar to those described elsewhere, where magnetic drums have een used for the storage of information controlling electro-echanical equipment.[3,4]

The accuracy obtained with such timing techniques depends pon the number of pulses counted and upon their spacing. The ossible timing error will always be equal to the interval between ming pulses, and the percentage error will be inversely propor-onal to the number of the count. In the above example an ror of 25% is possible, which in that application would be olerable. For other timing operations more stages of counting ould be used. Each additional binary stage halves the error ut adds one bit of storage capacity to the requirements for each hannel.

Fig. 12 also illustrates another general time-sharing principle. lthough the information on the highways is available at 10 kc/s, e counting operations are performed only during the timing ulses and the information in S1 and S2 need be presented only these times. At other times the same physical apparatus ay be used to perform other counting operations using the me pulse channels. Thus the logical counting elements may e time-shared by sampling, not only over the pulse channels ut over the functions required for each channel. Similarly hen stored information is changed infrequently—as in S1 and 2—stores having a much longer access time than 100 microsec ay be used. For example, a 1 500 microsec delay-line cir-ulating system capable of storing 15 × 80 bits of information ay be used to store 15 unrelated bits for each of 80 channels. the register apparatus it is convenient to use these 15 bits r information relating to the one connection to the calling ne using this channel. Thus four such stores may be used store in binary form the designation digits sent in over the coming line, in which each digit is stored in binary code on oincident bits in each of the four lines. Ten of the 15 positions ay be used in this way, while the remainder store instruction nd other information. In the supervisory apparatus it may be ore convenient to use the 15 bits for each channel on the same nction for each of 15 groups by allocating particular positions the cycle to each group. Using this technique, much of the gical apparatus used to control connections can be made ommon to the whole exchange.

Thus many of the control functions may be carried out using e pulse channels themselves to obtain a high degree of time aring. Other apparatus cannot be time-shared in this way. or example, if voice-frequency signalling is used for junction nd inter-register signalling, the information to be received in the ontrol apparatus is not conveyed merely by the presence and osence of the pulse channels, and the modulating signal must e recovered in order to receive the information. When expen-ve speech-guarded multi-frequency receivers are required, a onsiderable reduction in their numbers can be achieved by onnecting them to a line only when a signal is actually present be received. This economy may be achieved on incoming nction and manual-board calls by connecting a cheap wide-and receiver in the register apparatus to the incoming line. nly when this so-called detector indicates that a signal is present be received need the expensive receiver be connected to the hannel. The expensive receiver may be released and used on nother connection as soon as the signal has been recognized. his technique is particularly applicable when acknowledge gnals are used, since a channel may be prevented from being econnected to an expensive receiver until it is known that new nformation is being presented. A very small number of multi-oice-frequency receivers are then required, and the economy worth the extra switching and selecting. A similar tech-ique may be used for the supervisory apparatus where an xpensive voice-frequency receiver can be time-shared by the channels using cheap voice-frequency receivers as detectors on the junctions.

(5.2) Supervisory Control

In order to make the application of these time-sharing techniques more clear, a brief outline of the operation of the supervisory and register apparatus will be given. For local calls the supervisory apparatus must return ring tone to the calling party, ring the called party's bell, respond to the called party's answer condition and register the appropriate fee against the calling party, who also controls the release of the call.

Supervisory apparatus is provided for each group, so that on a call between two subscribers in different groups two sets of supervisory equipment will be involved. Some functions are more conveniently carried out on the incoming side and some on the outgoing side. Because the highways carry both-way traffic, the apparatus of any group may be dealing with incoming and outgoing functions on different calls. The signalling paths provided for the supervisory apparatus are shown in Fig. 13.

Each group of lines is connected to its supervisory apparatus over the group highways and also over a common lead used to control ringing. This common lead is connected by gates to a trigger—a cold-cathode thyratron—provided for each line, which, when operated, switches ringing current to the line. The trigger is reset by the negative half-cycles of ringing current, and is operated by applying the channel pulses used by the line over the common lead from a store in the supervisory apparatus. The answer condition on the outgoing line will cause the channel pulse to be generated continuously on the group's transmit highway, and this may be detected by a suitable timing operation and used to delete the channel pulse from the ringing store.

The answer condition is also indicated to the supervisory apparatus of the incoming group over a signalling path controlled by the within-group or group-to-group connection stores. Fig. 13 shows just the group-to-group signalling paths. On the incoming side the answer condition deletes the channel pulse from a store sending ring tone to the incoming line, and also initiates metering. The fee to be charged is stored using the pulse channel used by the incoming line. One store is used to indicate a unit-fee call, and upon receipt of the answer condition it sends a meter pulse over a common lead used for this purpose. This meter pulse is gated to the line, where it is used to operate the meter by means of a trigger circuit. By suitable gating the same common lead, line gates and triggers may be used for both ringing and metering by interleaving the meter pulses with the ringing.

The release signal is detected by timing the disappearance of the channel pulse from the incoming line's group transmit highway. The resulting indication is used to delete the channel pulse from all the connection memory stores, so releasing the connection.

This sequence illustrates the general way in which connections through the exchange may be controlled by apparatus which is common to large numbers of lines. Using the voice-frequency techniques already mentioned, junction calls may be handled by the same type of apparatus. With a junction signalling multiplex and ringing and metering paths connecting lines to group supervisory apparatus, and with signalling paths transmitting information between the supervisory apparatus of the two groups involved, any of the normal facilities required in a telephone network may be provided.

When the connection to the outgoing line is set up, the information relating to the subsequent control must be indicated to the supervisory apparatus and stored there, using the channel pulse or pulses selected for the final connection. Thus the marker not merely marks the outgoing lines but also indicates the required supervisory information, which is derived from a

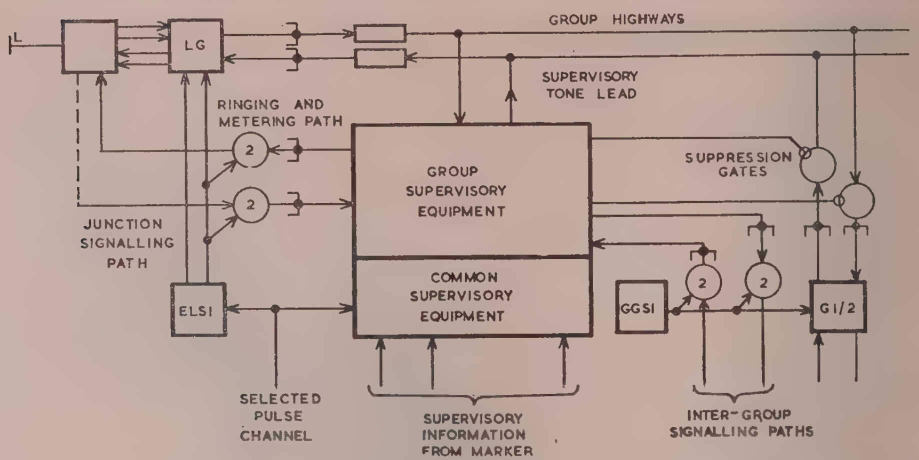


Fig. 13.—Group supervisory apparatus, showing the signalling paths used for inter-group connections.

L—Line.
LG—Line gates.
G—Inter-group gates.
ELS—External-line store.
GGS—Inter-group stores.

knowledge of the designation digits and of the class of service of the lines and parties involved.

(5.3) Register Control

(5.3.1) Receipt of Register Information.

When a register is connected to a calling line its class of service may be derived using the third signalling multiplex CSM shown in Fig. 11. The multiplex control lead of each line is connected to some combination of class-of-service pulses CSP1–5, which characterizes its class of service and causes the multiplex to generate the line-signalling pulses at times coincident with the CSP pulses to which it is connected. Thus each sequence of five pulses indicates in binary code the class of service of the line. When a line is selected its coded pulses are transmitted to the register apparatus, where the class of service is detected and indicated. Voice-frequency junction and key-sending manual-board class of service indications may be used to associate a voice-frequency detector with the connection in the register apparatus, and these and other classes of service may be stored using the selected pulse channel in other stores which are used for subsequent indication to the supervisory apparatus and for the control of the register operations.

The designation digits are received using either dial-break and other timing circuits or the voice-frequency detectors and multi-voice-frequency receivers. These digits are inserted, under the control of logical apparatus, into appropriate positions in the digit stores. These digits must be converted into signals which mark the suitable outgoing lines and which, if required, indicate to the supervisory apparatus the fee information. In addition, it may be necessary to translate the digits received into other digits used to further the progress of the call in other exchanges.

(5.3.2) Translator-Marker.

The general translation of digits into other signals can be carried out using a time-shared translator-marker which is switched on to the connection when sufficient digits have been received. Depending upon the prefixes and codes used in a national numbering scheme, the number of digits required to be received before a connection can be set up may vary. This number of digits may be determined from the first one or two digits in some elementary translation process carried out on the information stored as coded pulse channels.

The translator-marker staticizes the digits required to deri the signals which mark the outgoing line or lines and indicate th digits to be sent to the next exchange if any, the fee for the call ar any other relevant information. The marking of the outgoing lir or lines initiates the selection and setting-up process which leads the selection of an outgoing line and suitable pulse channels. Th outgoing line's class of service is derived in the manner alreac described and used in the register apparatus; for example, t control appropriate inter-register signalling which would b effected via an a.f. link connecting the register apparatus to th outgoing register highways. If voice-frequency inter-regist signalling is used, a voice-frequency detector is connected to th outgoing line and the digits are sent to the line using time-share voice-frequency senders.

The required supervisory information is derived from the tran lator-marker and from the incoming and outgoing lines' class of service. The incoming line's class of service is stored in th register apparatus by means of the channel pulse used for th incoming line-to-register connection, and will therefore have be staticized for indication to the supervisory apparatus.

The register apparatus will release itself from the connectic when the outgoing line connection is set up, unless the connectic is made over an outgoing junction, when the register could releas for example, on the receipt of ring tone. Also, when the registe releases, busy or n.u. tones may be reverted to the incoming lin using the incoming line-to-register pulse channel in the incomin line's group supervisory apparatus.

(5.3.3) Instruction Digits.

The register and supervisory apparatus and the translato marker are all general-purpose units which are used for all cal of all types and operate to provide a variety of facilities dependin upon the information transmitted to and stored in them. Th flexible use of these and other units may be further influenced b the increased use of instruction digits, which may be transmitte between registers in either direction using voice-frequency code similar to those used for digit sending. For example, the classe of service of the parties involved could be indicated to all th exchanges in a connection. This would enable all junctio circuits to be fully available for all classes of call and woul facilitate the application of other time-sharing techniques. Fo example, the same junction signal could be used for variou

urposes depending not only upon the state of a call but also pon the parties involved. This would enable the same voice-requency apparatus to be used for all classes of call. An ncreased use of instruction digits would also enable a greater use f alternative routing to be made. This would involve the ranslator-marker in making more than one attempt to set up a all over different routes, depending upon instruction information resented to it.

It would also be possible to centralize the more complex ranslating apparatus by indicating the exchange of origin of he call to a central translator-marker, which could then serve nany exchanges in an area and would operate in different ways lepending upon the exchange of origin and other information resented to it as instruction digits. Such a central control ould also use instruction digits to indicate information back to he exchange of origin for the subsequent control of the call. It vould then be possible to set up a connection by control appa-atus housed in an exchange remote from those through which he speech connection was established. These time-sharing tech-iques might find wide application in rural areas, where centrali-ation of the more complex apparatus would appear economic. Space does not permit the detailed development of these ideas, ut it is clear that time sharing need not be restricted to problems vithin any one exchange.

(6) DESIGN APPROACH

The operation of the switched-highways system has been lescribed by considering the functional aspect of its various arts. From the design point of view, the different types of pparatus are more conveniently distinguished by considering he methods which they use to handle the information. The line peech gates, group and register highways, a.f. links, and voice-requency senders, detectors and receivers all handle speech-requency signals, and their design and operation are to a great xtent determined by the speech and signal levels and by the ransmission performance required. Similarly the line-ringing nd metering apparatus is largely determined by the signals equired to operate the subscriber's bell and meter.

(6.1) Use of the Binary Code

As has been shown, all this apparatus may be controlled by he application of pulses to various gates and modulators. Also, ll the information relating to a connection is indicated to the ontrol using the presence and absence of signals on the line-ignalling paths, on the outputs of the voice-frequency detec-ors and receivers, and on the group and register highways. Thus all the information in the control uses binary elements, o that the detection of information involves merely the distinc-ion between the presence and absence of a signal. This implies hat the control can consist entirely of simple "and" and "or" ates together with binary stores such as triggers and pulse tores. This use of the binary code gives great flexibility in the rovision of facilities and service, since any sequence of operations nay be controlled using binary elements of this type.

Some of the control apparatus is used merely to convert ignals of one type to those of another. Thus the signalling nultiplexés convert d.c. signals separated in space to signals eparated in time, and staticizors convert pulse signals separated n time or space into d.c. signals separated in space. Standard nits of each type may be provided which are identical for all uch functions required in the exchange.

In the translator-marker, logical apparatus operates on informa-ion presented as d.c. signals separated in space. All the remain-ng control apparatus consists solely of pulse stores and gates quipped with output amplifiers appropriate to the units they upply. These stores and gates are time-shared equivalents of relays and contacts used in electro-mechanical systems, and can be subjected to the same logical design techniques. It has already been shown that the connection memory apparatus can be time-shared by coding the information so that each gate uses a different combination of stores. Similarly in the supervisory and register apparatus the information may be coded so that each possible state of a connection corresponds to a combination of stores in which its channel pulse is stored. Up to 15 different fees, for example, may be stored using the combinations of four stores in the supervisory apparatus, and a fifth may be used to indicate whether timed or untimed metering is required. These stores would normally be used for this purpose only on the incoming side of the connection and would be used for entirely different functions on the outgoing side.

Time sharing by coding has been applied widely in electro-mechanical systems, but from mechanical-design considerations the number of contacts which can be controlled by any one relay is less flexible than the number of gates which can be controlled by a single electronic store when the only factors involved are the power supplied from the store output amplifier and the cabling with which the amplifier is connected to the gates. Thus, using electronic techniques, time sharing by coding can be carried further and information can be handled in larger blocks.

(6.2) Component Operating Speed and Function Requirements

The underlying principle of design is to associate any item of apparatus with any function for as short a time and as infrequently as can be tolerated without loss of reliability or service, and many examples of its application have been given in which apparatus serves many different connections. Many operations which cause very different results on the connections are logically very similar, and it is possible for quite large blocks of apparatus to be time-shared by allocating them to different functions for different periods of time. As already indicated, many counting and timing operations are required, but since the repetition time required for these operations is considerable, comparatively few counters need be provided even in the largest exchanges. Similarly signalling paths, e.g. between the super-visory apparatuses on each side of a connection, may be time-shared by being allocated to different functions at different times. In these examples the sampling rate is matched to the time scale of the function to be performed.

It is also possible to match the speed of operation of the apparatus to the pulse durations required for various indications. Thus, whereas the highways and gates controlling speech pulses must use, say, 1·2 microsec pulses, this need not be the pulse duration used in the logical and storage apparatus if simple conversion equipment is used. Mercury delay lines used in the connection memory apparatus, for example, can each store pulses of, say, 0·4 microsec duration, and, by suitably gating, delaying and lengthening the pulses, can serve for three stores indicating their information using 1·2 microsec pulses.[6,7]

(6.3) Components

The paper has been restricted to a discussion of the principles and system techniques which can be applied using electronic techniques, and little has been said about the components and circuit techniques. In general, there are a number of alter-native techniques which can be used to perform any par-ticular function, and it is not easy to predict which will prove the most satisfactory. Many of the more promising tech-niques are still under development and no accurate estimate of the ultimate cost of the system can yet be made. There is good reason for believing, however, that an all-electronic exchange using the time-sharing techniques described in the paper will cost less than an equivalent electro-mechanical exchange.

(7) MAINTENANCE

The cost of maintenance also depends upon the components used, so that no accurate estimate of the reliability of the system can be given. It will be appreciated that the wide use of time-sharing techniques does not reduce their reliability. It does, however, introduce complexities which may make the detection of a fault from its service symptons a more difficult task. It will also cause each fault to affect the service of a greater number of lines. It is therefore desirable to provide automatic means of detecting and correcting faults. The use of check signals and error-detecting codes may find application in some equipment—e.g. in the translator-marker and setting-up apparatus—but, in general, some means of correcting faults must be found. This can be achieved using error-correcting codes or some technique which switches in spare equipment to perform the functions of a faulty unit. This does not necessarily mean duplication, since many like units, such as pulse stores, could be served by a single spare. This is not as unattractive as it may appear at first, since each spare unit is common to at least, say, 500 lines, and much may be made common to the whole exchange by time sharing. All the apparatus—including spares and spare-switching gates—could then be routined, for example, using a pulse channel reserved for this function. It is estimated that a routine of the whole exchange could be carried out in something under 30 sec, and if fault detection is used to switch in spares, the resulting service to subscribers would be at least adequate. Some warning of faults could also be obtained using marginal testing in which the voltage supplies to the exchange are varied during a routine. Both of these techniques should make an appreciable saving in maintenance costs since the number of urgent alarms would be greatly reduced.

(8) CONCLUSIONS

It has been shown that the principle of time-sharing apparatus, which has played such a large part in the development of past systems, can be applied to an even greater degree to electronic systems in the solution of all the basic problems which arise. Time sharing, by sampling, switching and coding has been the principle behind the design of the switched-highways system, which it is hoped has served to illustrate techniques that will find wide application in the future.

(9) ACKNOWLEDGMENTS

The author is indebted to many colleagues for assistance in the subject-matter and preparation of the paper. In particular, thanks are due to Mr. T. H. Flowers, whose many contributions to the art have made these developments possible, and to Mr. S. W. Broadhurst, for their many helpful suggestions. Acknowledgment is also made to the Engineer-in-Chief of the Post Office for permission to make use of the information contained in the paper.

(10) REFERENCES

(1) Flowers, T. H.: "Electronic Telephone Exchanges," *Proceedings I.E.E.*, Paper No. 1266, March, 1952 (**99**, Part I, p. 181).

(2) Lewis, W. D.: "Electronic Computers and Telephone Switching," *Proceedings of the Institute of Radio Engineers*, 1953, **41**, p. 1242.

(3) Baker, G. T.: British Patent Nos. 717687 and 717734–40.

(4) Mathaner, W. A., and Vaughan, H. E.: "DIAD—An Experimental Telephone Office," *Bell Laboratories Record*, 1954, **32**, p. 361.

(5) Flood, J. E., and Harris, L. R. F.: British Provisional Patent Application No. 16315, 1953.

(6) Newman, E. A., Clayden, D. O., and Wright, M. A. "The Mercury-Delay-Line Storage System of the Ace Pil Model Electronic Computer," *Proceedings I.E.E.*, Pape No. 1527 M, August, 1953 (**100**, Part II, p. 445).

(7) Harris, L. R. F., and French, J. A. T.: British Provisiona Patent Application No. 23887, 1953.

(8) Jacobaeus, C.: "A Study on Congestion in Link Systems, *Ericsson Technics*, No. 48, 1950.

(11) APPENDIX. TRAFFIC PROBLEMS IN THE SWITCHEI HIGHWAYS SYSTEM

(11.1) Connections between Lines in Different Groups

The basic traffic problem arising in the switched-highway system is that of finding the probability of not being able to mak a call between two free lines in different groups. A connectio between such lines can be set up only if there is at least or pulse channel which is free in both groups.

If N = Number of pulse channels.
A = Total traffic offered to each group.
P_x = Probability that x channels are busy in one group, GPx
P_y = Probability that y channels are busy in the othe group, GPy.
$P_{x,y}$ = Probability that there is no channel free in both GP and GPy when they have x and y channels bus respectively.
B = Grade of service (i.e. probability that the call is lost

$$B = \sum_{x=0}^{x=N} \sum_{y=0}^{y=N} P_x P_y P_{x,y} \quad . \quad . \quad . \quad ($$

B can be expressed in terms of the traffic offered to each grou if it is assumed that P_x and P_y are given by the normal Erlan distribution and that the busy channels in each group a randomly and independently distributed.

The Erlang distribution gives

$$P_x = A^x/x!(1 + A + \ldots + A^N/N!)$$

and
$$P_y = A^y/y!(1 + A + \ldots + A^N/N!)$$

$P_{x,y}$ is the probability that all the free channels in o group, say GPx, coincide with busy channels in the other grou GPy. If the busy channels in both groups are randomly di tributed over the channels, all arrangements of x and y a equally likely and $P_{x,y}$ is given by the ratio of the number arrangements of the $N - x$ free channels in GPx which can b made so that they all coincide with the y busy channels in GP (i.e. ${}_yC_{N-x}$), to the total number of possible arrangements of th $N - x$ free channels over all N channels (i.e. ${}_NC_{N-x}$).

Thus
$$P_{x,y} = x!y!/N!(x + y - N)!$$

and
$$B = \sum_{x=0}^{x=N} \sum_{y=0}^{y=N} A^{x+y}/N!(x + y - N)!(1 + A + \ldots + A^N/N!)$$

and if $x + y - N = z$

$$B = [A^N/N!(1 + A + \ldots + A^N/N!)^2] \sum_{x=0}^{x=N} \sum_{z=0}^{z=x} A^z/z! \quad . \quad ($$

This expression is the same as that derived for other blockir problems[8] and may be computed. Results for various values A and N are shown in Fig. 14.

(11.2) Discussion of Assumptions

The accuracy of expression (2) depends upon the validity the assumptions made. Since some calls are lost when not a

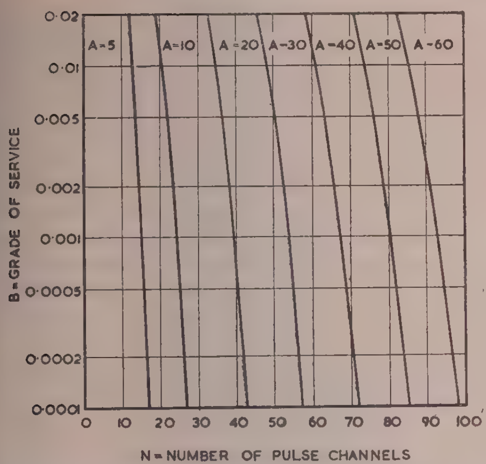


Fig. 14.—Grade of service for inter-group connections, assuming Erlang distribution and a random arrangement of busy channels, i.e. theoretical minimum traffic capacity.

A = Both-way traffic offered to each group.

he channels in a group are busy, the use of the Erlang distribution is not strictly correct, although with any normal grade of ervice, say one lost call in 500, the effect on the distribution s small. It can be shown that the Erlang distribution gives esults which underestimate the traffic-carrying capacity of such system. Thus if one group, GPx, is considered and p_x is the robability that a connection can be made to a free line in the roup when x channels are busy, the probability of losing a call s given by

$$B = \sum_{x=0}^{x=N} [1 - p_x]P_x = 1 - \sum_{x=0}^{x=N} p_x P_x \quad . \quad . \quad (3)$$

If calls originate at random—which is a reasonable assumption ince the number of lines in a group may greatly exceed the umber of channels—and the system is in statistical equilibrium, hen

$$p_x A P_x = (x+1)P_{x+1}$$

nd since

$$\sum_{x=0}^{x=N} P_x = 1$$

$$P_x = (p_0 p_1 \ldots p_{x-1})A^x/x![1 + p_0 A + \ldots + (p_0 p_1 \ldots p_{N-1})A^N/N!]$$

nd

$$B = 1 - \sum_{x=0}^{x=N} p_x (p_0 p_1 \ldots p_{x-1})A^x/x![1 + p_0 A + \ldots + (p_0 p_1 \ldots p_{N-1})A^N/N!] \quad . \quad (4)$$

If the Erlang distribution is assumed, the approximate grade f service is given by

$$1 - \sum_{x=0}^{x=N} p_x A^x/x!(1 + A + \ldots + A^N/N!) \quad . \quad . \quad (5)$$

This expression gives a pessimistic value of the grade of service if

$$U = \sum_{x=0}^{x=N} [p_x (p_0 p_1 \ldots p_{x-1})A^x/x!](1 + A + \ldots + A^N/N!) >$$

$$V = \sum_{x=0}^{x=N} (p_x A^x/x!)[1 + p_0 A + \ldots + (p_0 p_1 \ldots p_{N-1})A^N/N!]$$

Each of these expressions may be written as the sum of terms nvolving $A^k/(k-r)!r!$, for integral values of r from 0 to $k/2$.

The coefficient of $A^k/(k-r)!k!$ in U is $p_0 p_1 \ldots p_{(k-r)} + p_0, p_1 \ldots p_r$, and in V it is $p_r p_0 p_1 \ldots p_{k-r-1} + p_{k-r} p_0 \ldots p_{r-1}$, and the coefficient in U minus the coefficient in V is

$$p_0 p_1 \ldots p_{r-1}(p_r - p_{k-r})(1 - p_r \ldots p_{k-r-1})$$

This is positive for all values of k and r provided that $p_r > p_{k-r}$. Since the probability of losing a call must increase with k this condition is satisfied.

The assumption that the busy channels in a group are randomly arranged is true only if no within-group calls are made, if the two groups are entirely independent and if the channels are selected and released at random. In practice, perhaps only the last of these assumptions is valid. However, a random arrangement of the busy channels is by no means the most desirable arrangement from a traffic point of view. If it could be arranged that all the busy channels in one of the groups were the same as some of those busy in the other, calls would be lost only when all the channels were busy in one or other of the groups. This would give an approximate grade of service of

$$B = 2A^N/N!(1 + A + \ldots + A^N/N!) \quad . \quad . \quad . \quad (6)$$

which sets the most optimistic limit for the grade of service. The upper and lower values of traffic which may be carried by N channels are shown in Fig. 15, with a grade of service of 0·002.

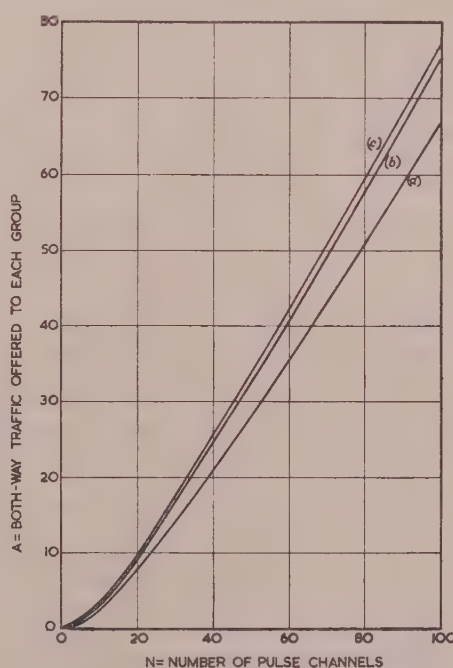


Fig. 15.—Traffic-carrying capacity of switched-highways system for inter-group connections and a grade of service of 0·002.

(a) Assuming Erlang distribution and random arrangement of busy channels, i.e. theoretical minimum.
(b) Theoretical maximum.
(c) Full availability.

It will be seen that an improvement of about 10% is possible if it can always be arranged that the same pulses are busy in both groups. Fig. 15 also shows the full-availability traffic for the same grade of service.

Although such complete dependence between groups cannot readily be achieved in practice, it would be possible to effect some improvement by selec ing channels in some preferred order and

by using channels which are already in use on other highways for other connections whenever possible. Tests with an artificial traffic analyser have indicated that some advantage is to be gained from such techniques, although the scale of tests did not indicate the amount of improvement possible.

(11.3) Other Traffic Problems

The traffic data given in Fig. 15 apply to connections between two lines in different groups carrying the same total traffic. They also apply to connections between calling lines and registers where the incoming register traffic is the same as the total traffic carried by a group. If a connection from a calling line can be made to a line in any of a number of groups, or if the size of the exchange requires more than one set of incoming register highways, the grade of service obtained on such calls will be improved and the traffic per group for a given grade of service will approach the full-availability traffic if a large number of alternative highways are suitable for the connection.

Such a choice of highways would exist if the outgoing line were a junction, since the lines of a junction route could be distributed over a number of groups. However, this advantage is then somewhat offset by the necessity to choose a channel which is also free on the outgoing register highways.

The only arrangement which is then suitable for computation is one in which the connection must be made to an outgoing line in one group and when only one set of highways can be used for the register connection. A channel must then be found which is free on three highways, two carrying the same traffic and one carrying less. Again, assuming an Erlang probability distribution and a random arrangement of pulses, this arrangement can be computed, and it can be shown, for example, that with 80 channels the estimated grade of service is 0·003 if 50 and 30 erlangs are carried by the group and outgoing register highways, respectively. These figures indicate that an adequate grade of service can be given if there is a choice of group or outgoing register highways and if preferred selecting techniques are used.

The worst grade of service would be given for within-group calls when it was necessary to use a pair of adjacent channels. The effect of this restriction is difficult to calculate and would be considerably influenced by using preferred selecting techniques such that whenever possible channels were selected which would leave a maximum number of pairs of channels free in each group.

The within-group traffic originating in each group is equal to the total traffic carried by the group divided by the number of groups, so that the total within-group traffic in the exchange is approximately equal to the total traffic carried by a group of lines. It would therefore be possible to use a single common group of a.f. links. For each connection, two pulse channels would be required which are free on the group's highway and on the a.f. link highways. The grade of service for such an arrangement would compare unfavourably with the group-to-group grade of service, but it would be much improved by reducing the within-group traffic by making group-to-group connections to p.b.x. lines and junction routes whenever possible.

(11.4) Conclusions

Sufficient information has been derived to show that the traffic-carrying capacity of the switched-highways system would be adequate. This capacity is improved if various preferred selecting arrangements are used, when the traffic per group might approach the traffic offered to a full-availability group for a given grade of service. More accurate traffic data are hardly required at the present stage of development.

DISCUSSION BEFORE THE RADIO AND TELECOMMUNICATION SECTION, 9TH MAY, 1956

Mr. G. C. Hartley: The author implies that the current developments in switching are just a slight extension of the old principles, and that there is really nothing new in the system. That is misleading, and I do not think it will help us in this new field. It must be accepted that some of the ideas are basically new.

The author has undoubtedly given a wide definition of time sharing. I think it is rather misleading and likely to lead to confusion to refer to the age-old habit of using common registers or even common groups of junctions as time sharing in the sense that we need to discuss it, which, I believe, should be confined to operations which we conduct with cyclical patterns—and that field alone is wide enough. Even here there are some very major distinctions between pulse multiplex in the handling of the speech transmission itself, and time sharing in the logical operations of handling numerical data, etc.

The author has already indicated that the main divisions of the switching problems are the switching-network aspect, which provides the actual conversational path and the long-term memories to maintain it, the selection process, which is a transient operation and should include the release or disconnection, and supervision, which also involves long-term memories but a slower time scale and narrow bandwidth.

The author has given, first, in general terms a two-wire both-way connection. But later he makes it quite clear that he is pinning his faith on a four-wire switching system. I am fairly certain he must be aware that techniques are now available for time-division coding which are capable of providing both-way two-wire communication, and not only that, but also of conveying over the channel the entire energy content of the speech.

Such systems impose far more onerous demands on the power-handling capacity of the gates, but I am not at all certain that this difficulty will not be removed in due course. If it is, the two-way version will have some major economic attractions. Similarly the author's coded combination use of memories is neat and ingenious, although it involves certain costs in coincident gates, but it is probably an indication of a feeling that the memory itself is undoubtedly a little expensive.

There again, I believe that new memories (and possibly one of the most promising is the square-loop ferrite type of memory) can, with their low cost, transform the situation and enable extremely simple controls to be used.

Finally, on the speech-path issue, I would urge that we must not, in surveying the wonderful era of miniaturized facilities which can be given by time-division switching, lose sight of the extreme importance of ensuring that the transmission standards are adequately maintained.

When the author talks about gating a thousand lines on to one highway, I believe that he is taking on a fairly formidable job.

On the selection aspect, we want to keep a clear distinction between the selection of the line and the actual control of the set-up of what I call the switching network.

Supervision is immensely fascinating. There are two extreme approaches. One is the acceptance of the idea that memories for speech are to be held throughout the conversation and that the same channels can be used. The other is that it should be entirely divorced from speech and that an independent data processing type of network should be used to handle supervision.

The author sometimes seems just a little dangerously half-way between the two and not getting the best of either approach.

Mr. D. A. Barron: The author's definition of time sharing seems to be sound, but it must be realized that, because the definition necessarily covers the general case wherein apparatus

is used for different connections at different times, every telephone system in operation in the world is a time-sharing system. In the Strowger system used in the United Kingdom, for example, it is well known that the amount of apparatus particular to any subscriber is very small. The essential problem is to determine what sort of time sharing is to be used, and how it shall be arranged. A practicable result will normally be a compromise between first cost, on the one hand, and complexity, service risk, and maintenance difficulty, on the other. Further, time sharing can affect the traffic overload capacity of a system, and I should be glad of the author's comments on this point.

With regard to maintenance, it seems certain that self-testing arrangements will have to be 'built into' any successful electronic system, but I am not very satisfied about the reference to marginal testing. There is always difficulty in determining suitable margins, in knowing for how long the equipment would continue to function satisfactorily even if the margin were not achieved, and in ensuring that any maintenance attention resulting from the marginal testing will not of itself lead to additional faults. For electro-mechanical equipment, we are moving away from limit testing in favour of functional testing.

A primary feature of electronic exchange systems is the high speed of operation, and it is of interest from the development aspect to consider where, initially, the best use could be made of this feature. High-speed sending from a subscriber to an electronic local exchange will give a high setting-up speed for local connections. It must be remembered, however, that, for a considerable time, a large part of any existing national switching system will continue to be electro-mechanical. It seems, therefore, that one of the more important early applications of electronic switching techniques would be in the field of trunk switching, so that, in conditions of fully automatic trunk operation, the setting-up time for multi-link switched calls could be reduced.

Lastly, I should like to mention a point in connection with fully automatic international telephone operation. An originating international register translator has to examine the 'country' code digits to select the appropriate outgoing route to the objective country. For charging purposes, it has also to determine the rate, and for countries comprising several charge areas it may be necessary for the register to examine not only the 'country' code, but also perhaps three or more digits of the national number of the required subscriber. With electro-mechanical equipment, such an increase in the number of digits to be examined results in a serious increase in the cost of the register translators. Can the author comment on the degree of improvement in this respect which may result from the use of electronic techniques?

Dr. J. H. Mitchell: I recall that when Mr. Flowers presented a paper* on this subject before The Institution I was among those who were severe critics of his proposals for an electronic telephone exchange. I am therefore particularly pleased to see how much progress has been made in this development over the last four years. The author's system, in contrast with the earlier proposals, shows a welcome economy of apparatus, albeit at the cost of rather more flexible and elaborate pulse-producing equipment.

It is difficult to make other than general remarks in the absence of detailed circuit proposals, but I would have thought it preferable so to choose pulse widths and timings as to make use of magnetostrictive delay lines, perhaps in conjunction with transistors, in place of the mercury delay lines proposed in the paper. I feel that the advantages of solid over mercury delay lines are such as to outweigh any slight circuit disadvantages involved. This may require the use of smaller switching groups, with a consequent disadvantage in trunking arrangements, but loss of economy in this respect should be very slight. As an example of a switching system making use of rather longer pulse durations in this way, I would refer to a Swedish paper* describing a system which uses 10 microsec pulses produced by hydrogen-filled cold-cathode tubes. With this system it is possible to envisage an exchange which contains no hot-cathode tubes at all. While this may not be the best approach, it is one which has many attractions, and I wonder whether the author can give any idea of the number of hot-cathode tubes per subscriber in his system?

The principle of time sharing, while valuable as a means of reducing the volume of apparatus, should not be pressed too far in the interests of ease of maintaining a good standard of service. I see, for instance, that the connection circuits are based on common storage; failure of one of these could cause faulty or double connections in a large group of circuits. Although a fault of this type can easily be detected and cleared, it has an extremely high nuisance value.

With regard to costs, statements have been made in the United States that electronic exchanges will cost much less than electro-mechanical exchanges. I trust this news has not unduly coloured the opinion expressed in Section 6.3, as American methods and costs bear little resemblance to the practice in this country. I think that cost predictions are deceptive—at least until a fully developed prototype exists.

Mr. T. H. Flowers: In the switched-highways system every circuit terminating on the exchange has to have two terminals—the transmit modulator and the receive gate and demodulator. This is fairly expensive equipment on a per-line basis. But when this has been provided, extension to any number of lines on the exchange involves merely one stage of highways switching. This switching involves relatively little apparatus and cabling. Because it is multiplex, the cabling for one highway is equivalent to cabling for perhaps 100 a.f. circuits. Hence the switched-highways system shows to the greatest advantage in the large exchange.

For the smaller exchanges—100 up to perhaps 2000 lines—it is economical to use a simpler type of multiplex, often called a two-wire or both-way gate. This involves an amount of equipment on a per-line basis which is just about equal to what has been disclosed for space-multiplex switches using diodes. With the two types of units, therefore, all sizes of exchanges with all sorts of configurations can be covered in an economical manner.

With regard to the registration of meter pulses, it is undesirable from several points of view that the present mechanical meters be used. One possibility is to use ferrite cores for number storage. Ten per line would accommodate numbers up to 1023. Time-shared common equipment would add meter pulses to the records by reading the existing numbers and re-writing with the addition of one unit. The same common equipment could be used to transfer the reading, at perhaps daily intervals, to magnetic tape, which, by various check tests, would prevent a serious loss of information owing to faults. Not a little of the attraction of such a system would be that it would provide a fully-mechanized accounting system.

Mr. N. C. Smart: In contrast with the views of Mr. Hartley or Dr. Mitchell, it is my belief that the elements shown in the paper make an entirely satisfactory and most attractive automatic exchange system.

In particular, I believe that the development from time sharing by switching to time sharing by sampling and coding is the next logical step in the development of automatic exchanges. I think that it will rank with some of those other historic develop-

* Flowers, T. H.: 'Electronic Telephone Exchanges', *Proceedings I.E.E.*, Paper No. 1266, March, 1952 (99, Part I, p. 181).

* Svala, G., and Jacob, W.: *Teknisk Tidskrift*, 1955, 85, p. 807.

ments, such as that of the step-by-step selector and crossbar selector.

The spur that has urged work in this field is the desire to produce an exchange, which, while having all the facilities of present exchanges and being comparable in cost, will operate with substantially less maintenance. It should also occupy considerably less space. As a result of our work in this field I am confident that the time-division multiplex switched-highways system will meet these demands.

It is surprising, and gives me some gratification, that after some eight years of independent work we have arrived at almost the same conclusions as the author.

I should like to end with two notes of warning. First, this is a research project so far, and the stages that lie ahead in development, engineering and manufacture are likely to be difficult. There may be some unseen pitfalls in the way, and certainly there are some well-recognized difficulties.

Secondly, the production of such a system will entail great changes in our manufacturing organizations. The time of transition will be difficult, but I am sure that the industry can make a success of a new system of this type.

Mr. C. E. Calveley: The system described in the paper requires a hybrid for each subscriber's line, which is expensive, and it is interesting to learn from Mr. Flowers of a two-wire electronic exchange system which avoids this expense. However, perhaps the author would state whether it is possible to put part of the system he has described in some cabinet towards the subscriber's termination, using d.c. switching technique from that point into the highways exchange and so offset to some extent the cost of a hybrid for each subscriber by the saving in copper on the subscriber's cable. If so, what would be the approximate ratios of the subscriber's terminations to exchange pairs at the remote switching point?

In the Introduction, the author refers to expensive multi-voice-frequency signalling receivers, and as we do not have this form of signalling in this country, could the author give further details of the system visualized? If automatic alternative routing of trunk calls is economically justified, a faster signalling and switching is required than we have at present. If we have the faster signalling system for the numerical information which has to be transmitted over the line, the time occupied would appear to be so small that there need be no great expense in associating voice-frequency receivers with the register instead of time sharing from all registers to a common multi-voice-frequency signalling receiver.

If we reach the stage visualized by other speakers, in which the exchange switching equipment requires little maintenance, equal attention will have to be paid to the essential ancillary services, lest we reach the stage where more maintenance effort will have to be given to the auxiliary services than the main switching equipment. One of the most important of these services is, of course, power, and perhaps the author would state what power supplies would be required for this equipment and whether its reliability will be equal to that visualized for the switching equipment.

Mr. Barron has already compared the overload capacity of the step-by-step equipment with the system described by the author. I would like to take this a stage further by asking whether any steps could be taken to avoid a complete lock-up in the exchange when a large subscriber's cable goes faulty. In Section 7 there is a reference to the possibility of switching in spare equipment to perform the functions of a faulty unit. If there is only one marker for the whole exchange, is it possible to switch any spare equipment to replace a faulty unit in the marker? Would it not be essential to provide a standby marker with change-over arrangements? Would it be necessary to go a stage further and put in some space division with the markers? If this is desirable, is it necessary further to complicate the circuit arrangements by providing for marker lock out?

Mr. E. P. G. Wright: There is little doubt that time sharing is generally accepted as a basis for switching, but its attractiveness depends on the availability of suitable components. In Section 7 it is stated that, 'The cost of maintenance also depends upon the components used, so that no accurate estimate of the reliability of the system can be given.' But surely the practicability of electronic switching depends on an accurate estimate of the stability and reliability of the components. It should not be necessary to prescribe the meticulous care lavished on submarine repeaters, but something better than ordinary commercial components may be required.

The possibility of rapid routining is valuable, perhaps essential, for a pulse system because of the difficulty of locating intermittent faults except by repeated tests. There is an important difference between finding that something is faulty and ascertaining the cause. This suggests that electronic system design must cater for simple maintenance as a factor of primary importance. The proposed spreading of junction and p.b.x. lines over different groups may well improve trunking efficiency and consequently reduce initial costs, but this arrangement would complicate, and therefore increase, annual charges for maintenance.

The author wisely emphasized that error-indicating codes can assist maintenance, but it seems necessary to plan not only fault indication but also the means to assist in locating and replacing faulty units. Efforts in this direction with digital computers have not been very successful.

A schedule of component quantities would form a valuable extension to the paper, and permit comparison to be made with other time-sharing proposals.

It is to be hoped that electronic systems will be planned to reduce maintenance charges, and it would be desirable to have a paper specifically directed to the subject of employing electronic means for assisting in the maintenance of electronic systems.

Mr. L. J. Murray: It appears that a determining factor in the design of a time-division multiplex system is the limiting, by present sampling techniques, of the number of pulse channels in one highway to 80 or 100. Previously, this has been overcome by the connection of highways in tandem selecting stages with a.f. links between them.

The switched-highways system overcomes the difficulty by dividing the exchange into groups and switching between them by a multiplicity of inter-group gates.

The author shows an electro-mechanical equivalent of his switched-highways system which over-simplifies the switching arrangement between groups. He should really have shown the inter-group gates as double-bank switches.

Mr. Flowers has assured us that the cost of the switched highways system is very satisfactory for the larger exchanges, but the number of inter-group gates tends to rise with the size of the exchange. For instance, five groups means ten gates, ten groups mean 45 gates; 20 groups mean 190 gates. As the number of groups is doubled, the gates increase fourfold. The cost of that important item of switching equipment therefore tends to rise. Is the cost offset by the sharing of the common equipment by more lines?

The author states that there is little loss of economy if the number of groups does not exceed 50, but I submit that there is an important minority which will require more groups, e.g. the central London type of exchange with a high originating traffic and a large percentage of junction traffic. Such exchanges would seem to require 60–70 groups.

The author states that, if the number of groups is very large, some economy may be achieved by interconnecting the high

ways in stages, using multiplexed equivalents of well-known electro-mechanical arrangements. Will the author state what these are?

The switched-highways system raises the problem of within-group connections. It seems that the author is making a big assumption in saying that the within-group traffic originating in each group is the total traffic carried by the group divided by the number of groups. This may be true as an average over all groups, but in any one group it may vary widely; and for that reason, in handling within-group traffic, any arrangement which restricts the choice of pulse channels should be avoided. It would therefore appear that perhaps the best scheme is that of Fig. 3, which uses a group of a.f. links.

Dr. J. E. Flood: It is unfortunate that shortage of space has prevented the author from discussing the difficulties of transmitting time-division-multiplex with a large number of channels over substantial lengths of cable without introducing inter-channel crosstalk. In order to achieve adequate crosstalk attenuation, it would appear necessary to use complicated equalization of the pulses after transmission or some form of predistortion of the pulses before transmission, e.g. the curbed pulse system developed by Moss and Parks.* Either method is likely to be expensive.

It is therefore an attractive idea to arrange each highway so that there is no possibility of adjacent pulses occurring. In

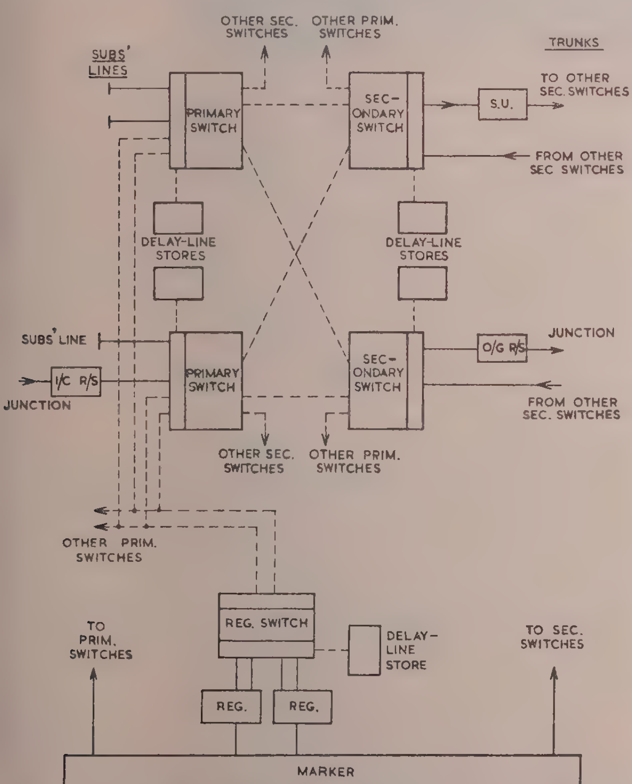


Fig. A.—10000-line electronic exchange using delay-line storage.
——— Audio-frequency and direct current.
— — — Pulses.

Section 2.3 it is pointed out that if an audio-frequency path is inserted with additional gating it is not necessary to make all pulse channels available for all the possible routes through the exchange. Use of one audio-frequency link for every connection also eliminates the difficulties of within-group connections discussed in Section 2.5. Another advantage is that the audio-frequency link can contain supervisory apparatus. If a common-control type of data-processing system is used for supervision it need not be synchronized with the transmission system, thus enabling an entirely different technique to be used. If these ideas are followed, they lead to an arrangement which is a time-division analogue of the link system used in crossbar switching frames. We have made some study of this arrangement,* and a possible trunking scheme for a 10000-line exchange is shown in Fig. A. The trunks between frames use a.f. transmission and contain the supervisory units. The links between pairs of switches use pulses whose timings are fixed and can be chosen to be non-adjacent. Adjacent pulses may occur on the links to the register switches, but these do not carry conversations. The transmission problem through the exchange is thus much simpler than in the switched-highways system, and there is greater freedom in arranging the supervisory apparatus.

Mr. A. J. Bayliss: The three main developments which have made the type of telephone exchange described in the paper possible are, first, the efficient use of the multiplex channels by joining sufficient lines to each multiplex unit so that all the channels are fully loaded at the busiest time of the day, secondly, the provision of an economical source of pulses for operating the gates in the subscribers' line circuits (see Fig. 7), and thirdly, the use of the switched-highways system itself, which enables transmission to be made through an electronic exchange with only one modulation and one demodulation process.

We have been working, quite independently, along similar lines to the author, and have built a model telephone exchange which employs these principles. There is a channel pulse store corresponding to ELS1 of Fig. 8, consisting of a number of magnetostriction delay-line units, each of which may hold one-hundred $\frac{1}{2}$ microsec pulses. It serves the subscribers' line units. Control of the exchange is also largely by means of magnetostriction delay lines. This model exchange has been in operation for some months and is giving satisfactory service.

The saving in space which can be achieved by using electronic switching techniques can be very considerable. Fig. B shows a model of a building which normally holds up to 800 lines of mechanical switching and which is shown laid out as an electronic exchange. Each of the ten racks of apparatus on each side represents a multiplex unit serving a group of, say, 500 subscribers. Down the centre are racks of apparatus for junctions, register translators, and common control equipment. The saving in space is quite evident.

Mr. P. W. Ward: Before we built our model there was an attitude that the trunking theory, on paper, was better than would be expected. Then we found that the author had come to the same conclusions, and I can support him in that there seems to be no stage at which the scheme becomes blocked. The traffic which can be passed between groups on a common channel appears to be only 10% less than that given by Erlang's formula for a single group.

We agree that the storage can be quite efficient and the memory item is not particularly expensive. The cost does not increase alarmingly with a large exchange with many different connections between the groups, since to some extent the principle of coding can be used.

The author in his trunking system and use of pulses exemplifies a slight difference between the scheme he has proposed and the scheme which we are using. This involves a problem which we have not had occasion to tackle. I have applied the formula

* Moss, S. H., and Parks, G. H.: 'Pulse Communication on Lines', *Journal I.E.E.* 1947, 94, Part IIIA, p. 503.

* Flood, J. E.: 'A Model Electronic Telephone Exchange' *Siemens Brothers Engineering Bulletin*, No. 273, May, 1956.

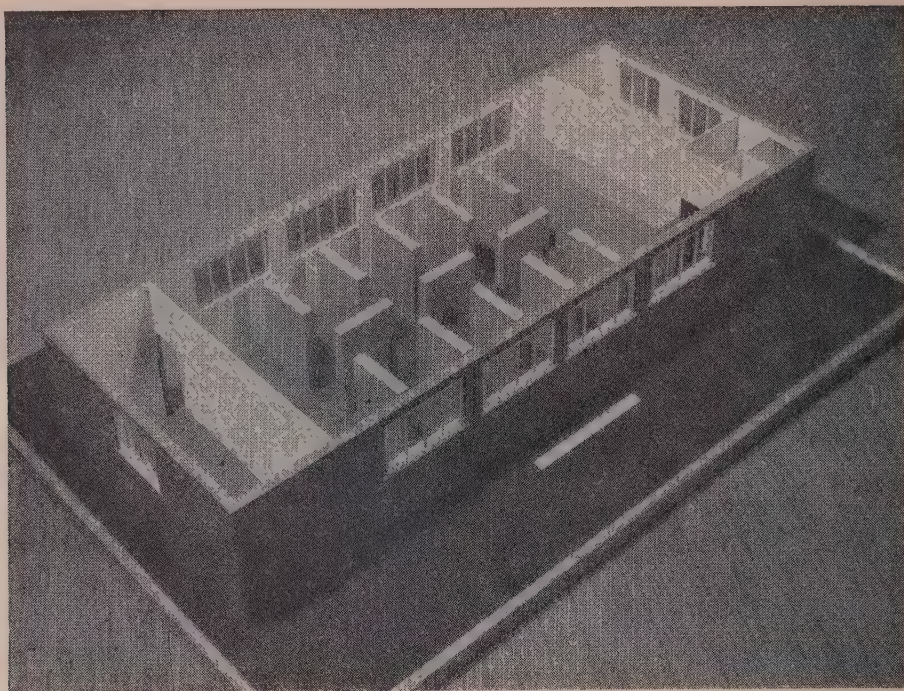


Fig. B

put forward some time ago by Dr. Flood to the delay which relates to one or two pulse periods, T or $2T$. It appears that this would tend to be a source of crosstalk. Are there any problems with that, because crosstalk does not appear to be the problem it might seem, and it is a pity that there should be what is apparently a possible source of crosstalk in the system?

Mr. R. Syski: I wonder why the author did not calculate the indicated summations in eqn. (2). When this is done the result is

$$\sum_{x=0}^{N}\sum_{z=0}^{x}\frac{A^z}{z!}=\sum_{i=0}^{N}(N-i+1)\frac{A}{i!}$$

Hence

$$B = E_N(A)[N+1-A+AE_N(A)]$$

where $E_N(A)$ is the Erlang loss formula.

This formula is much easier to calculate than the long formula given by the author. Incidentally, $B = \partial AE_N(A)/\partial A$.

The expression for B is equivalent to eqn. (20) in Reference 8. The author's treatment is then mathematically equivalent to Jacobäeus's treatment of the common control circuits (e.g. registers) when they receive the same traffic as each group of trunks (cf. also Section 3.2.2 of Rodenburg's paper,* with $\beta k = 1$). Although the switched-highways problem is more involved than the problem of common control circuits, the latter may perhaps serve as a first approximation.

Did the author consider explicit expressions for p_x in Section 11.2? Perhaps Rohde and Störmer's† theory of passage probabilities could be used, after the necessary modifications. It is useful to note that for the Erlang distribution all the p_x's are unity except for $x = N$ which vanishes. This throws light on the approximation involved in eqn. (5).

The author mentions in Section 1.4 that 'time-sharing principles . . . introduce no change in the approach to the basic problem of designing a telephone system using a minimum of apparatus for a given service'. Does this refer to the traffic calculations also?

For the exact calculations of the lines in the exchange, the interdependence of stages and groups should be taken into account. This means that the whole exchange should be considered as a unit and not in the separate parts. This can be achieved by joint study of the switching algebra and congestion theory. So far little has been done along these lines. In a recent paper* the connection was established between the switching algebra and traffic calculations. This approach would be very useful here.

Mr. G. W. Thompson: The parallel switched-highways system described by the author comes under the category of a blocking type of network, in that coincident pulse channels must be available before a telephone path can be set up between subscribers in two separate groups, or adjacent pulse channels must be available for a connection between subscribers in the same group. Quite recently a study of non-blocking systems has been made in America† and the conclusion has been reached that new forms of electronic control apparatus can possibly make a non-blocking network an economic proposition for equivalent grades of service.

With conventional types of electro-mechanical exchange systems and also with the switched-highways scheme as described, a speech path is set up on the completion of receiving the dialled information, and it remains set up until the calling subscriber replaces his handset. Statistics show, however, that the amount of actual speech information transmitted during this period is very small indeed. However, if one could visualize an electronic control arrangement which could set up a potential path in a time which was infinitesimal compared to the actual speech pulse transmission time, it is not beyond the bounds of possibility

* RODENBURG, N.: *Communication News*, 1953, 13, p. 69.
† ROHDE, K., and STORMER, H.: *Mitteilungs Blatt für Mathematische Statistik*, 1953, 5, p. 185.

* LEE, C. Y.: 'Analysis of Switching Networks', *Bell System Technical Journal*, 1955, 34, p. 1287.
† CLOS, C.: 'A Study of Non-Blocking Switching Networks', *Bell System Technical Journal*, 1953, 32, p. 406.

that a system could be designed which would only set up the desired path each time intelligence was there to be transmitted. Such a system would, of course, require individual register devices per line—a facility offered by the magnetic drum, for example—so that the routing information would always be available each time speech intelligence had to be transmitted.

Mr. S. W. Broadhurst (*communicated*): The degree to which time-sharing techniques may be applied to the design of a telephone switching system is dependent on the components available to the designer. Hitherto, in electro-mechanical systems, time sharing has been almost exclusively confined to the use of switching and coding techniques, and the limit of its application has been determined by the mechanical durability rather than by the speed of the components used. Sampling techniques have been impracticable, and since also each unit of storage has required a relatively expensive static device, common controls such as registers have each been provided with individual storage and logical sequence elements. Such apparatus is provided economically only if its holding time can be minimized and its utilization increased by time sharing by switching.

When high-speed sampling techniques are available in association with cheap electronic storage, a different approach is possible. If information is stored in binary form as pulse patterns in a cheap storage, and if large groups of such patterns can be processed by common apparatus sampling the sources of information, a redundancy of storage can often effect economies in the switching. Thus, although in an 80-channel system of the type described fewer than 80 registers may be needed for the traffic, it is probably cheaper and easier to provide one register per channel than to introduce another stage of switching. Similarly, it is cheaper to provide every trunk with the facilities of a universal cord circuit rather than to segregate the various classes of traffic.

Cheap storage may make it preferable to provide service facilities, such as observation and traffic recording, as built-in features of every channel, rather than to use auxiliary equipment time-shared by switching.

A redundancy in the speech channels is possible if and when devices such as high-speed transistors become available, and this could clearly lead to a system in which traffic blocking could not occur.

It is perhaps worth pointing out that an increase in the degree of time sharing, given reliable switching components, decreases the fault liability of the system since fewer components are used, and although the vulnerability of the system becomes greater and some duplication of the logical apparatus may become desirable, the use of delay-line storage simplifies the routine testing, since any pulse train is suitable for testing the whole of the storage.

Mr. W. T. Duerdoth (*communicated*): A large switched-highways exchange suitable for 10 000 lines will be of such a size that the lengths of highways interconnecting some of the group equipment to the inter-highway switch may need to be 100 ft or more. These long highways cause distortion in pulse shape and interchannel crosstalk. The highways have high characteristic impedances and the pulses have d.c. and a.f. components, so that a 75-ohm screened cable cannot be used for the highway without some intermediate equipment. This equipment can take the following form.

The pulses are fed to a short-circuited delay line with a time delay equal to half the pulse length. The delay line is driven from a resistance equal to its characteristic impedance, and the resulting voltage takes the form of a positive pulse immediately followed by a similar negative pulse. This double pulse contains neither d.c. nor a.f. components and can be passed via a transformer on to a 75-ohm coaxial cable. At the receiving end similar equipment can be used and a unidirectional pulse selected. Nevertheless, with cable lengths of 100 ft the crosstalk is still excessive, but this may be removed by adjustment of the resistances terminating the short-circuited delay lines.* Adjustment of the resistances causes multiple reflections to occur which take the form of progressively decreasing replicas of the original pulse, and crosstalk may be balanced by these reflections. This is a cheap method of removal of crosstalk and would seem to be applicable to considerably longer cable lengths.

The possibility of transmitting pulses over a mile of cable gives the switched-highways exchange a facility which would enable some of the group equipment to be located at a number of points near the periphery of the exchange area. Highways would then connect the remote groups to a centre which would contain the remaining groups and the inter-highway switch. Such an arrangement should show considerable saving in line plant.

THE AUTHOR'S REPLY TO THE ABOVE DISCUSSION

Mr. L. R. F. Harris (*in reply*): The idea of developing an all-electronic telephone exchange system has become generally accepted during the last few years, and the discussion showed a gratifying measure of agreement on the sort of system techniques which seem most promising at present. Many of these enable apparatus to be time-shared, and it seems a useful idea to relate the new techniques with those of past systems using the principle of time sharing. I am therefore surprised that Mr. Hartley should feel that the restriction of the term 'time sharing' to sampling techniques would be helpful, particularly as the introduction of electronic techniques considerably influences the degree to which apparatus can be time shared by coding and switching.

In the paper I refrained from discussing circuit techniques in any detail except to illustrate more general principles, and although various contributors have made reference to particular components, I feel it would be unwise to state a preference for one type of component over another. The more promising components—transistors, ferrites, crystal rectifiers, regenerative stores, etc.—together with their associated circuits, are still under development for telephone and other purposes, and to give component quantities or cost estimates at this stage would be unrealistic. With present components it is estimated that the switched highways system would be cheap and reliable enough to warrant further effort on its development. The saving in space illustrated by Mr. Bayliss provides a most convincing confirmation of this view.

Some apprehension is felt by contributors about the difficulties of maintaining a system dependent upon common and somewhat complicated apparatus. The magnitude of this problem will depend upon the components used and must be taken into account in the design of any system. Various duplicating, sparing, routining and checking techniques are mentioned in the paper and discussion, and each of these will probably play an appropriate part in the final solution. The experience gained with computers should help to estimate the merits of marginal testing and other techniques. The components used will also determine the type of power supplies, and I agree with Mr. Calverley that these and other ancilliary services must be adequately reliable.

I endorse Mr. Hartley's point concerning adequate transmission standards. The problem of connecting 1 000 lines to one set of highways is certainly formidable, but, in my opinion, soluble nevertheless. The both-way gates mentioned by him and Mr. Flowers offer a most attractive means of making the connections,

* British Patent Application No. 21327, 1954.

and I expect improved components to enable them to handle rather narrower pulses than at present. They can then be applied directly to the larger exchanges involving highways switching without uneconomic reduction in the number of pulse channels.

While agreeing with Mr. Ward that there is a crosstalk problem, it is worth noting that adjacent pulses are never both delayed by $2T$. Also, I believe that Dr. Flood exaggerates the difficulties of transmitting speech samples over cable, and I am grateful to Mr. Duerdoth for his contribution on this point. His comments, together with the both-way gates, are relevant to the question raised by Mr. Calverley as to whether parts of the system can be put out into the local network. By dispersing the exchange a valuable economy in line plant could be achieved, although it might require some modification of the techniques used at the switching centre. The number of subscriber's terminations for each set of highways depends upon the traffic carried and the possible number of pulse channels.

I am grateful to Mr. Syski for pointing out the simplification of the expression in eqn. (2). I have not considered explicit expressions for p_x in Section 11.2 and am interested in his comments on this. I agree that, for exact calculations, the exchange must be considered as a whole. However, at present it seems somewhat premature to go beyond showing that the traffic-carrying capacity is adequate. Mr. Ward provides confirmation that this is so, and answers a query raised by Mr. Barron. With regard to Mr. Murray's comments, within-group connections are an unfortunate complication to the traffic problem no matter which technique is used to make them. However, the effect can be minimized by making group-to-group calls whenever possible. This is facilitated by spreading the p.b.x. and junction lines over the groups, and with reference to Mr. Wright's contribution, I cannot see why doing this would materially affect maintenance charges. The blocking due to cable faults mentioned by Mr. Calverley can be overcome by incorporating suitable 'time out' and 'parking' facilities.

Various alternative trunking arrangements which make use of the same basic elements are discussed in the paper. Which will prove to be the most satisfactory will depend upon many factors, including the effects on local line plant transmission, signalling, flexibility and so on. That suggested by Dr. Flood is of interest, but apart from the transmission point already mentioned, I suggest that there is no great advantage in using different techniques for supervision. The place for separate data processing seems to be in the registration of meter pulses, as suggested by Mr. Flowers. Elsewhere the effective marriage of the speech and supervisory control apparatus, so that the same channels and techniques can be used for both, seems to offer considerable advantage, particularly where voice-frequency signalling is used. Here Mr. Broadhurst's contribution is very relevant, and I particularly agree with his point that time sharing by sampling can frequently effect economies in switching. His comments are also pertinent to Mr. Hartley's discussion of supervision. As he and Mr. Ward point out, storage can be efficient and economic, and from the switching point of view the interpolation scheme put forward by Mr. Thompson sounds feasible. However, it seems probable that the saving in pulse channels would be outweighed by the cost of detecting when to make connections through the exchange.

The interpolation system, mentioned in the paper, by which voice-frequency receivers can be time shared by switching, is applicable to any voice-frequency signalling system, and the advisability of its use can only be determined by considering each application on its merits. Shortage of space prevents a more detailed discussion of Mr. Calverley's reference to this, but with economic switching and storage techniques there seems to be some economy to be had for both inter-register and junction signalling, particularly where large groups and multi-voice-frequency receivers are involved.

In answer to Mr. Murray's queries, the number of inter-group gates required for each group is equal to about half the number of groups. Even with a 70-group exchange this number is small compared with the number of line gates, and in any event the slight rise in the cost of these gates is offset by the higher degree of time sharing which can be effected in the control with more groups. The number of gates can be reduced by interconnecting the group highways, using more than one stage of highway switching. A paper* by Charles Clos is relevant to this problem. This paper is also relevant to Mr. Thompson's comments on non-blocking systems, which—as Mr. Broadhurst points out—can be approached if a redundancy in the number of speech channels can be tolerated.

I agree with Mr. Barron that the speed of electronic switching would be of benefit to trunk switching. So far as the international register-translator is concerned, the increase in the number of digits should not present any great difficulties. I suggest that the cost would be roughly proportional to the number of live codes needed, but the great advantage to be gained by electronic techniques in this respect is the degree to which the translator can be time shared.

Finally, I agree with the warnings given by Mr. Smart, and share his confidence that the difficulties in developing and making a new system can be overcome.

* Clos, C.: 'A Study of Non-Blocking Switching Networks', *Bell System Technical Journal*, 1953, 32, p. 406.

521.396.822

Paper No. 2147 R
Nov. 1956

AN INVESTIGATION OF ATMOSPHERIC RADIO NOISE AT VERY LOW FREQUENCIES

By F. HORNER, M.Sc., Associate Member, and J. HARWOOD, M.A., Ph.D.

(*The paper was first received* 16*th March, and in revised form* 23*rd May*, 1956.)

SUMMARY

A description is given of a technique for investigating the charac-
eristics of atmospheric noise, and of the type of information obtained
t very low frequencies. The results quoted are typical of those which
ave been obtained in southern England during a long period of record-
ng, but a discussion of the statistics of all the data is not included.

Atmospheric noise in a bandwidth of 300 c/s at frequencies in the
ange 10–35 kc/s is intermediate in character between fluctuation noise
nd discrete impulses. For example, the variation of the average
oltage with bandwidth is similar to that for fluctuation noise, but the
oise also contains peaks of high amplitude. The r.m.s. value of the
nvelope is of the order of five times the average value (compared with
·1 for fluctuation noise), which illustrates the impulsive nature of the
oise and shows that neither parameter by itself can provide a satis-
actory description of the noise. The long-term trends, however, have
een expressed in terms of the average field strength.

A more detailed description of the noise is given in terms of the
mplitude distribution of the peaks in the envelope and the amplitude
robability distribution of the envelope itself. Either of these distribu-
ions can be expressed empirically in terms of two parameters, and
an be interrelated at the higher voltage levels where the impulses have
onsistent shape determined by the characteristics of the receiver. At
ne location the amplitude of the noise may vary between wide limits
vith time, frequency and bandwidth, but the parameters which can be
sed to describe the noise structure are comparatively invariant.

To supplement the amplitude data, information is required on the
ime sequence of the voltage changes, and it has been found that the
arger impulses are randomly distributed in time. There is a tendency
or the smaller impulses to occur in groups, in accordance with the
nowledge that a single atmospheric often consists of a number of
uccessive discharges, but it is doubtful whether the slight departure
rom randomness would have a significant effect on the severity of the
nterference with the usual types of radio service.

(1) INTRODUCTION

In most measurements of atmospheric noise a receiving system as been used to record a particular noise parameter, such as he average or r.m.s. voltage, or the strength of signal required o provide a specified quality of radiocommunication. The arly work on this subject was surveyed in 1947, and an extensive ibliography prepared.[1] Since that time, measurements of noise t low and very low frequencies have been made in several ountries[2–6] and there has been a growing realization that one arameter is not always sufficient to describe the noise adequately or practical purposes; for instance, two samples of noise may ave the same average value but yet be quite different in structure nd in their interfering effects on a given system. Some means f describing the structure of the noise is therefore needed, referably by way of easily measurable parameters.

The paper describes a technique used to study the level and tructure of the noise at frequencies in the range 10–40 kc/s, and ives typical examples of the results obtained. The measure-nents have been confined to very low frequencies to avoid interference from radio stations, and have been facilitated by the relatively high noise level at these frequencies. The techniques are, however, applicable to higher frequencies in locations where the noise level is high compared with the general level of station transmissions and man-made noise.

(2) OBJECTIVES

An atmospheric noise field at an aerial is a random fluctuation with components at all radio frequencies.[1] At the output of a receiver or measuring apparatus the noise will have been modified by frequency selection, amplification and possibly by detection and amplitude limitation. The term 'atmospheric noise' may refer to the original field, to the output from the receiver or to a voltage at some intermediate part of the receiver. In seeking a description of the noise for general use it is desirable to eliminate as far as possible its dependence on the particular characteristics of the measuring apparatus. A measurement of the entire spectrum of the noise field is impracticable, however, owing to interference from station transmissions. Furthermore, the user will generally prefer to have information relating to his particular small band of frequencies to avoid the necessity for performing his own selection on the data. A frequency selection must therefore be made, but with this exception an attempt has been made to describe the flucutations in the noise field by allowing for the receiver gain and by minimizing or allowing for the effects of amplitude limitation.

The bandwidths used in this type of measurement should be as large as possible without leading to station interference, partly to obtain more information and partly because attempts to suppress noise interference by amplitude limitation in operational equipment may be made in an early stage of a receiver with low selectivity. The noise in the chosen bandwidth may be regarded as an oscillation of the centre frequency of the pass-band, modulated in amplitude and phase. The point of view adopted is that the interference caused by the noise is related mainly to the amplitude modulation, and that the phase modulation is of secondary importance: only the noise envelope has therefore been considered.

The basic problem is to describe the noise envelope in terms of parameters from which the interference to radio services may be deduced, bearing in mind the widely diverse types of modulation in present and possible future communication systems. The most elementary descriptions indicate only the general level of the noise, e.g. the average and the r.m.s. voltage, and do not enable the fluctuations in the envelope to be described. A comparison of the average and r.m.s. voltages will indicate whether the fluctuations are large or small, but a more detailed description is required, in terms of either the variation of voltage with time or its Fourier transform, the frequency spectrum. The time-variation presentation is thought to be more readily applicable to practical problems and has been used in the paper.

The variation of the noise envelope with time may be expressed in terms of the amplitude probability distribution, together with information about the time-sequence of events. Several workers

Written contributions on papers published without being read at meetings are nvited for consideration with a view to publication.
The paper is an official communication from the Radio Research Station, Department of Scientific and Industrial Research.

have investigated noise by statistical methods[7,8] and have suggested the amplitude probability distribution as a means for partially describing the noise. Some success has been achieved in relating the distribution empirically with the interference to certain types of communication system, but the time-sequence will also be important in some problems. An illustration may be given by reference to Fig. 1, which shows four envelope wave-

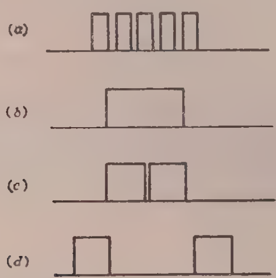


Fig. 1.—Idealized pulse sequences of noise.

forms consisting of pulses of uniform height. All have the same amplitude probability distribution, but (*a*) and (*b*), for example, may have quite different interfering properties, depending on the type of signal on which they are superimposed. A probability distribution of pulse widths would differentiate between (*a*) and (*b*) but would also differentiate between (*b*) and (*c*), which would often have similar interfering properties owing to the shortness of the break in (*c*). On the other hand, (*d*) has the same pulse durations as (*c*) but, with the greater separation, might have different interfering properties. Thus, we need to know, not only the distribution of the widths of the pulses, but also that of the intervals between them. The problem becomes more complex with pulses of different amplitudes or with a continuously varying voltage.

The description of the noise envelope in statistical terms is therefore a complex problem, particularly if a uniform system suitable for any type of noise is required. The problem has been approached by studying the noise to decide on the most essential characteristics and then attempting to describe these statistically. It is found that the main features are quite different at low and high frequencies, and different methods of description are appropriate; at the low frequencies discussed here the noise consists largely of a series of pulses whose shape is determined mainly by the bandwidth of the receiver, if this has a value such as is used in typical radio services. Amplitude probability distributions are given in a form appropriate to any type of waveform, but consideration of the impulsive nature of the noise facilitates the description of the time variations.

(3) NATURE OF THE NOISE AT VERY LOW FREQUENCIES

Atmospheric noise was received on a vertical aerial and passed through a tuned amplifier. Fig. 2 shows samples of noise obtained at 10 kc/s by displaying the output of the amplifier as a vertical deflection of a cathode-ray tube beam and photographing this on a film moving horizontally across the trace. Records are shown for two film speeds and the time scales are shown above the records. Of each pair, the upper was produced by atmospheric noise and the lower by diode fluctuation noise of about the same average voltage. It is evident that the character of the atmospheric noise differs appreciably from that of the fluctuation noise; records (*a*) and (*c*) show the occurrence of a number of impulses whose heights greatly exceed any which occur in records (*b*) and (*d*). The fast film of atmospheric noise shows also the characteristic shape of the discrete impulses; the shape

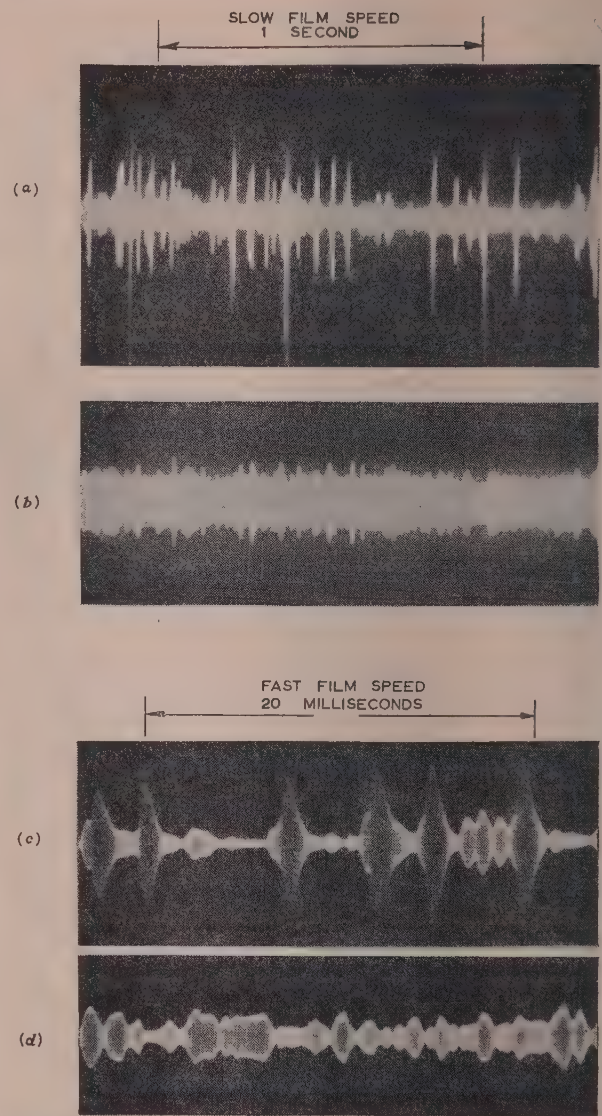


Fig. 2.—Samples of noise recorded at same average level on a radio receiver at 10 kc/s, with bandwidth of 300 c/s.
(*a*) and (*c*) Atmospheric noise.
(*b*) and (*d*) Fluctuation noise.

was very closely that which would be produced by a short pulse at the input to the tuned amplifier.

To describe atmospheric noise of the type depicted, the envelope of the amplifier output was selected either visually on a film or electronically with linear detectors, and statistical measures of certain characteristics were obtained.

(4) MEASURED AND DERIVED NOISE CHARACTERISTICS

The following characteristics of the noise envelope were either measured directly or derived from the measurements:

(*a*) Average voltage.
(*b*) Amplitude distribution of voltage peaks.
(*c*) Amplitude probability distributions of the voltage.
(*d*) R.M.S. voltage.
(*e*) Time distribution of voltage peaks.

These characteristics are expressed in terms of voltage, since they were measured at the output of the amplifier, and the dis-

cussion is in similar terms. All the final results, however, are given in terms of the equivalent field strength at the aerial, by ntroducing the appropriate instrumental factors. The frequencies used were 10, 23 and 33 kc/s, and the effects of varying he receiver bandwidth were investigated. The techniques of neasurement are described below.

(4.1) Average Voltage

The output of the tuned amplifier was applied to a detector ircuit in which the envelope was selected, and the average voltage lerived from an *RC* network having charge and discharge time-onstants of 8 sec. Records of the average voltage were obtained continuously on a pen recorder throughout the experiments and vere used as an index of the general magnitude of the noise under arious conditions of time, frequency and bandwidth. Smoothing over periods longer than 8 sec was carried out visually to the xtent required by the analysis; for example, the mean over a few ninutes was used in analysis of the structure, and over an hour n determining diurnal variations.

(4.2) Amplitude Distribution of Voltage Peaks

The impulsive nature of the noise suggests a measurement of he rate of occurrence of voltage peaks and the distribution of heir heights. A limiting circuit was used to select the portion f the envelope lying above a certain threshold voltage, and the umber of times per second that the envelope rose above the hreshold was counted (see Fig. 3), This rate will be called the

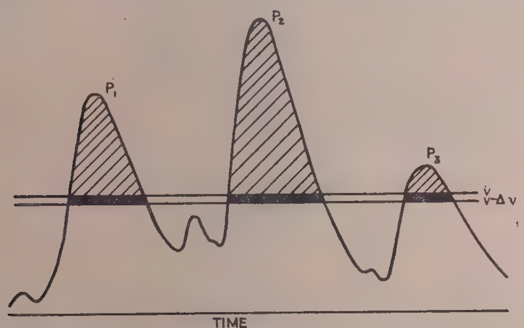


Fig. 3.—Diagram illustrating analysis technique.

Deductions from average level:

$\bar{V}_v$ = Average level above v (proportional to shaded area).

τ_v = Time spent above v = (solid area)/Δv = $-\Delta \bar{V}_v/\Delta v$.

Deductions from pulse rate:

If pulse shape is known, amplitudes P_1, P_2 and P_3 allow determination of shaded rea $\propto (\bar{V}_v)$ or width of pulses at $v(\infty \tau_v)$.

pulse rate', since the measurement has its main significance when pplied to the discrete impulses in the noise envelope. The hreshold was varied in ten steps at 5 min intervals, and a curve vas plotted of the pulse rate against the threshold; from this urve the distribution of the peaks themselves could be derived y differentiation.

(4.3) Amplitude Probability Distribution of Voltage

To obtain information about the amplitude distribution of the nvelope, the average voltage was measured of the portion of the nvelope appearing above selected thresholds.

These average values $\bar{V}_v$ were plotted as a function of the hreshold, v. By differentiation, the cumulative amplitude robability distribution [the probability, Q(v), that the envelope oltage is higher than v] could be obtained (Fig. 3), since

$$\mathrm{Q}(v) = \frac{d\bar{V}_v}{dv} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The probability density P(v) could be obtained by further differentiation, since

$$\mathrm{P}(v) = \frac{d[\mathrm{Q}(v)]}{dv} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The range of threshold voltages over which the above technique can be applied is limited by the difficulty of measuring the small average voltages which are obtained when only the portion of the waveform above a high threshold is selected. It has been found, however, that the shape of the larger peaks is nearly always that which would be obtained by applying a very short impulse to the receiver. Since this shape is known, a knowledge of the distribution of the heights of the peaks (see Section 4.2) enables the probability distribution to be calculated. For example, with three similar tuned circuits in cascade, the pulse shape is given by

$$v = \tfrac{1}{4}\hat{v}(\alpha t)^2 \varepsilon^{2-\alpha t} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where v = Amplitude at time t.
$\hat{v}$ = Maximum amplitude.
α = 6·15 times the overall bandwidth between 3 dB points.

By a graphical process the durations of such a pulse at different levels relative to the peak level may be found, and the recorded peak amplitude distributions can be converted to cumulative amplitude probability distributions.

There is normally a range of amplitudes for which both the above techniques are valid, and in this range they provide a check on each other.

(4.4) R.M.S. Voltage

It is evident that the r.m.s. voltage can be derived from the amplitude probability distribution P(r), since

$$v^2_{r.m.s.} = \int_0^\infty v^2 \mathrm{P}(v) dv \quad . \quad . \quad . \quad . \quad . \quad (4)$$

In practice, however, the value was obtained more directly from the plot of $\bar{V}_v$ against threshold v, since it can be shown that

$$v^2_{r.m.s.} = 2\int_0^\infty \bar{V}_v dv \quad . \quad . \quad . \quad . \quad . \quad (5)$$

The upper limit was actually taken at a finite level which was rarely exceeded by atmospheric impulses. The higher amplitude ranges provided a substantial contribution to the r.m.s. voltage, and it was clear that a quoted value would depend markedly on the chosen upper limit. The measured upper limit would sometimes be the overload level of the receiver, but allowance was then made for the clipping of the larger peaks.

(4.5) Time Distributions

Information on the time distributions of the peaks was derived mainly from analysis of film records of the noise envelope. Some of the data were obtained from a 10 millisec sweep on a cathode-ray oscillograph, and others from continuous records of the type shown in Fig. 2.

(4.6) Note on the Validity of the Technique

It will be observed that the success of the technique depends on the statistics of the noise remaining stationary over periods up to an hour. The consistency of the results obtained after exluding periods when gross changes in the average level occurred indicates that the assumptions made were valid.

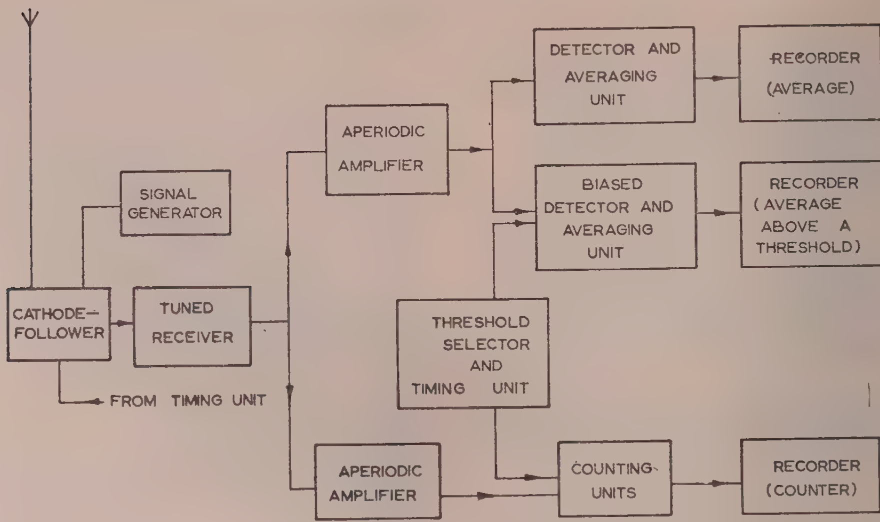


Fig. 4.—Block schematic of equipment.

(5) RECORDING EQUIPMENT

(5.1) Aerial and Receiver

A block diagram of the whole equipment is shown in Fig. 4. A vertical aerial was located about 70 ft from the recording equipment on a clear site and connected by screened concentric cable to a cathode-follower unit containing relay switches. These were used to disconnect the aerial for the setting of recorder zeros and to inject a signal from a standard signal generator to determine the calibration factors of the equipment. The calibrations were carried out automatically at predetermined intervals. The output of the cathode-follower unit was applied to a receiver with three tuned stages of approximately equal bandwidth. The overall bandwidth was normally 300 c/s between the 3 dB points of the response curve. To obtain convenient voltage levels for measurement and to supply many measuring circuits, the receiver was followed by aperiodic amplifiers, having conventional power output stages.

(5.2) Measuring Equipment

The average level of the envelope was measured in a unit of two similar channels. One channel provided a continuous record of the average level, while the other recorded the average above various threshold voltages. Each channel consisted of a linear diode detector, to whose anode the noise was applied, and an *RC* load to select the envelope of the applied noise. A threshold bias voltage was applied to the diode of one channel. Following the load was a series *RC* averaging circuit with charge and discharge time-constants of about 8 sec. An amplifier and recording milliammeter were used to record the direct voltage across the capacitor in the averaging circuit. A diagram of the unit is shown in Fig. 5.

The pulse rate was measured by applying the envelope waveform, after threshold selection and shaping, to a ratemeter followed by a pen recorder. Provision was also made for measurement of low counting rates on an electro-mechanical register.

Automatic operation of the equipment was achieved by means of a switching and timing unit. An electric clock motor was used to operate a set of relays and a uniselector at 5 min intervals. The unit changed the threshold voltages applied to the measuring units, put timing-marks on the pen records and operated the

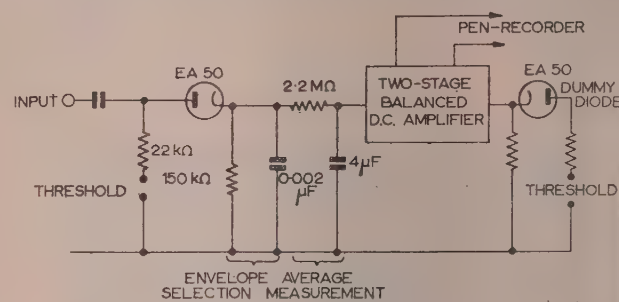


Fig. 5.—Circuit of unit measuring average voltage.

calibrating relays at the input cathode-follower unit. To obtain the sequence of threshold voltages, a d.c. voltage-divider was connected to the contacts of the uniselector.

(5.3) Form of the Pen Records

Fig. 6 shows a typical set of pen records obtained in a bandwidth of 300 c/s at 10 kc/s. The top recording shows how the average level of the envelope (with no applied threshold voltage) varied over a period of about three hours. This record was used as a reference with which to compare readings from the two lower recordings.

The middle record shows the average level above a series of decreasing thresholds, the last of which is zero; the vertical lines show when the changes of threshold occurred. A sequence of changes on the records was completed in one hour, and included an interruption in the reception of atmospheric noise during which zeros and sometimes calibration signals were recorded.

The bottom record shows the pulse rate above the series of decreasing thresholds. Presentation on linear scales of rate from about 0·1 to 100 per second was achieved by automatic range-changing of the ratemeter. This would initially operate on a range on which full-scale reading corresponded to one impulse per second; as the applied threshold decreased, the rate would exceed this value, and the range would change to give a full-scale deflection at 10 per second, and subsequently 100 per second. Fig. 6 shows records on the last two ranges only, th

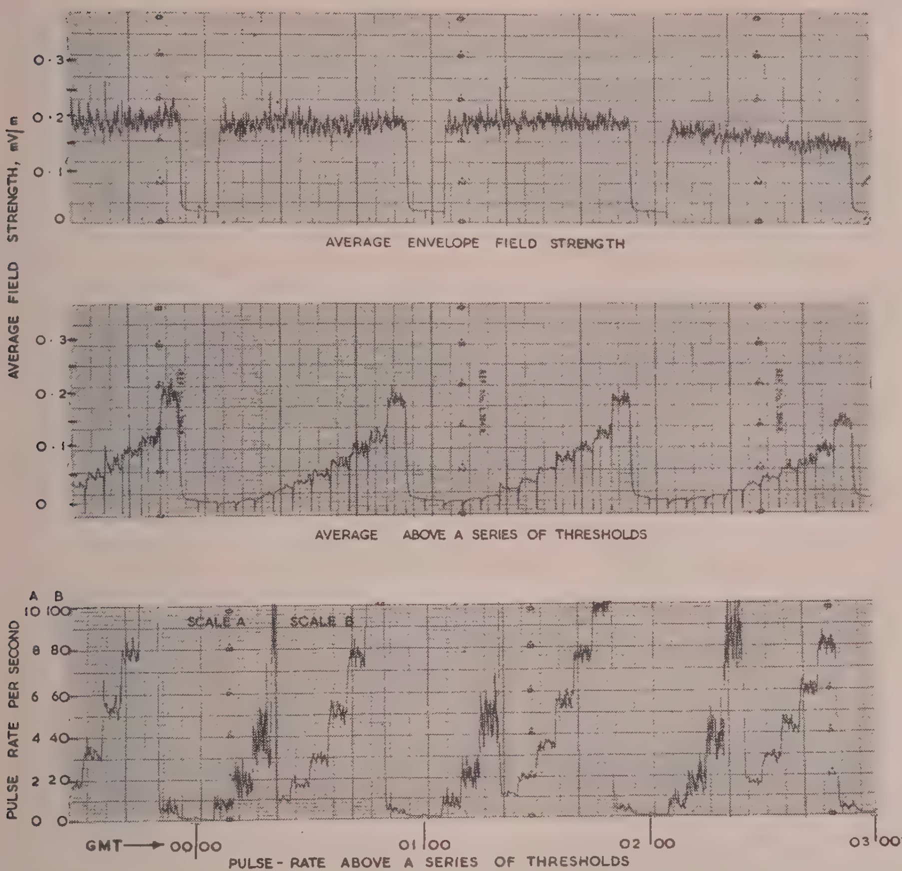


Fig. 6.—Typical pen records at 10 kc/s; 22nd August, 1955.

minimum rate being too high for operation on the lowest range. The pulse rate above zero threshold is of no interest, and therefore the last interval of the sequence was used to obtain an additional reading at a high threshold.

The calibration of the system was carried out with a standard c.w. signal generator at frequent intevals to obtain the gain factors of the various units and to check their linearity. A number of test experiments were also carried out with a pulse generator and with a source of fluctuation noise. Further checks were made by comparing measurements on photographic records of the noise with those on the pen records produced simultaneously by the equipment.

(6) RESULTS

(6.1) General

A comprehensive survey of all the results obtained is not attempted in this paper, but a few results are quoted to illustrate the method of analysis. The measurements were made in southern England.

Data were taken from records such as those shown in Fig. 6 by estimating visually the average level over each 5 min period. The trace showed fluctuations with a period of about 8 sec (the time-constant of the system) which were usually less than $\pm 10\%$ irrespective of time of day or seasons, except for a slightly greater scatter during the morning and a more 'spiky' type of trace during local storms. During a sequence the average level might exhibit a slow random change or part of the normal diurnal change; the random change was seldom greater than 20%, but near sunrise and sunset the change in an hour could be much greater.

To approach statistically stationary conditions in the analysed sequences, those involving changes of level greater than about $\pm 15\%$ (and later $\pm 10\%$) were ignored. No significant loss of data was occasioned by the selection. In the selected sequences a correction was applied to take account of the variation of average field strength when the other traces were being analysed. The correction was based on the assumption that for small changes of average field strength the amplitude distribution maintained the same form with proportionate magnification or reduction of the scale.

In the following description of the results the variations of average field strength with time, frequency and bandwidth are first considered, and then the structure is described from measurements involving the thresholds.

(6.2) Average of the Envelope

Most measurements of the average of the envelope were made at frequencies of about 10 and 25 kc/s; in addition, enough were made at 30–35 kc/s to give an indication of the magnitude and diurnal changes of level at these frequencies. The average for each hour (hourly value) and the median of the hourly values for corresponding hours on all the days of a month were obtained. Typical results are plotted in Fig. 7. At 10 kc/s the

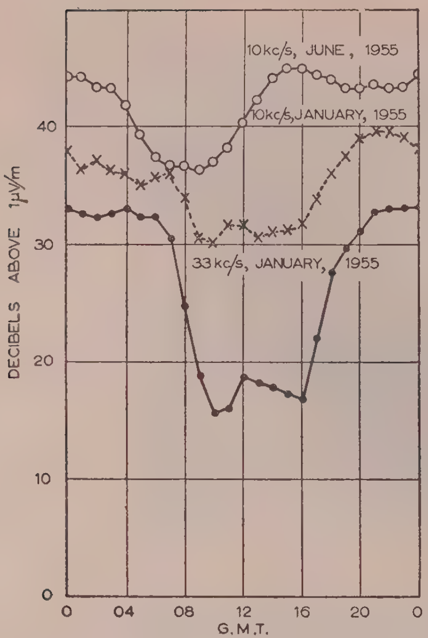


Fig. 7.—Typical diurnal variations of average field strength.

upper and lower decile hourly values were, respectively, about 5 dB above and below the median.

The curve in Fig. 7 showing the diurnal variation at 10 kc/s in January is representative of the general form of all the diurnal curves during times of low thunderstorm activity, with a lower level in the day than at night. Results during the summer months showed a high afternoon level, often exceeding the night-time values. At any time of day the average noise decreased with increasing frequency; the lowest level of the day occurred in the morning, and the overall diurnal change was greater the higher the frequency.

Although the general afternoon noise level was higher during summer than in winter, the effect of local storms on the average noise was not great. During the hours 1400–1600 in June, 1955, the mean field-strength for eight days when storms were reported within 500 km was 200 μV/m, while the mean for all days was 190 μV/m, and for three days when the nearest storm was 1 400 km away it was 165 μV/m. On one day with intense and widespread local storms the level reached 400 μV/m for a short period. Hence, it appears that in southern England the main contribution to the average noise level at very low frequencies is derived, in general, from distant storms.

The majority of experiments were carried out with a receiver bandwidth of about 300 c/s, but the variations in the characteristics with bandwidth were also studied. A 25 kc/s receiver was used with a single tuned circuit whose effective bandwidth could be varied by means of a feedback circuit in the range 25–500 c/s. During afternoons in early summer it was found that for bandwidths (f_b) up to about 200 c/s the average field strength was proportional to $f_b^{0\cdot5}$, as it would be for fluctuation noise, while at wider bandwidths the variation with bandwidth was rather slower. Tests were also made with a 10 kc/s receiver with three tuned stages, the overall bandwidth of which could be varied between 200 c/s and 2 kc/s. On two occasions (afternoons in April and September) the average was found to vary as $f_b^{0\cdot33}$ and $f_b^{0\cdot24}$ over this range.

The results are in accordance with the observation that the noise becomes more impulsive at the wider bandwidths, but even at 300 c/s, where the noise has a distinctly impulsive appearance (see Fig. 2) there is sufficient interference between successive atmospherics for the average field strength to be dependent on bandwidth.

(6.3) R.M.S. Value of the Envelope

The r.m.s. value of the noise envelope may be deduced from the curve of average against threshold, as described in Section 4.4, additional information about the high-voltage levels being derived from the pulse rates. In contrast with the average value, the r.m.s. is critically dependent on the high-amplitude parts of the waveform and the samples of noise considered must be of sufficient duration to include the largest peaks. From the experimental data it has been concluded that samples should be at least one minute long for this purpose. Analysis of a few such samples, in a bandwidth of 300 c/s at 25 kc/s, has given values in the range 4–8 for the ratio of the r.m.s. to the average field strength, compared with 1·13 for fluctuation noise. Over periods of a few seconds, however, ratios as low as 2 have been recorded, so the r.m.s. field strength is a more variable quantity than the average.

The pulse-rate measurements show that in a period of a minute some peaks of amplitude more than a hundred times the average voltage will occur. In operational equipment it is likely that amplitude limitation will occur on the larger peaks and the ratio of r.m.s. to average value will thereby be reduced. As the extent of the amplitude limitation will depend on the particular equipment the values have been given for a hypothetical receiver with no limiting action.

(6.4) Noise Structure

(6.4.1) Ampltitude Distribution of Peaks.

Information about the distribution of heights of peaks is given in the form of plots of the number of peaks per second (pulse rate) exceeding a given field strength against the field strength, both on logarithmic scales. A typical plot is shown in Fig. 8(a). The plots are usually found to be nearly linear from a field strength two or three times the average up to one hundred times the average or even more, at which level the pulse rate is of the order of one per second. The pulse-height distribution may therefore be represented by a power law

$$N = AE^{-p} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Where N = number of peaks per second exceeding the field strength E.

A and p are constants in this equation, but may vary with time, frequency and bandwidth. In practice, A, which is an index of the average field strength, varies considerably with time and frequency but p does not vary between wide limits with either variable, in the frequency range 10–35 kc/s. It lies between 1 and 2, and a small diurnal trend in the variations within these limits has been found. Eqn. (6) has no firm theoretical basis, but it may be noted that a uniform spatial distribution of equal lightning flashes would give a power law with $p = 2$, assuming inverse distance attenuation. This suggests that the variations in p might be related to the movements of storm centres and that it might be possible to establish a quantitative relationship if more were known about the storms.

The deductions of amplitude probability distributions at the higher levels were made from curves of the type shown in Fig. 8(a), as described in Section 4.3. Analysis of a number of special observations indicated that linear extrapolation of these curves to higher field strengths was valid, and this was done where necessary. It was assumed, however, that even the high-intensity

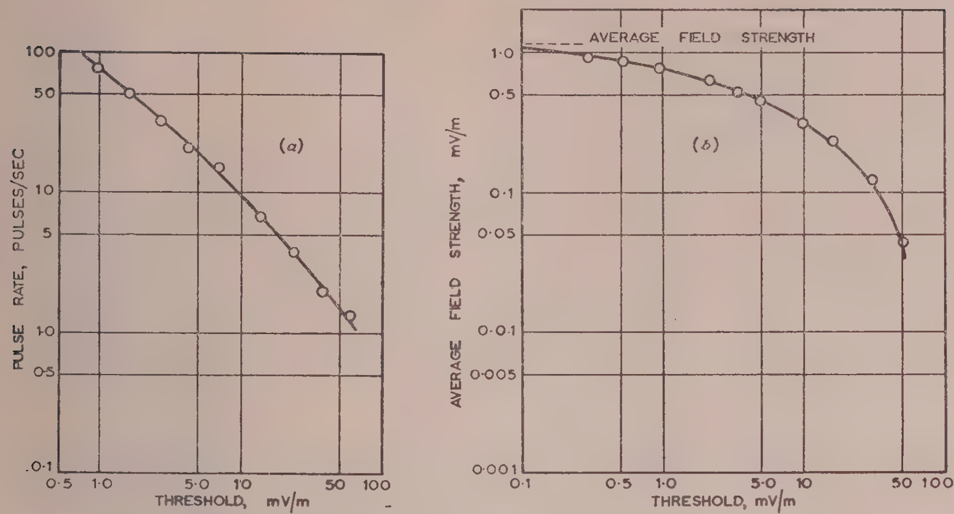


Fig. 8.—Pulse rate and average field strength at 10 kc/s; 1400–1500 G.M.T., 14th July, 1955.
(a) Pulse rate above threshold. (b) Average field strength above threshold.

peaks would not be of great practical importance if their rate of occurrence was much less than one per minute. Moreover, a finite upper limit to the amplitudes of the peaks must be assumed in order that the noise power shall be finite, and the amplitudes of the pulses occurring at the rate of one per minute appear to be about the correct order for the upper limit.

(6.4.2) **Amplitude Probability Distribution of the Envelope.**

The main information about the amplitude distribution of the noise envelope has been obtained from curves showing the variation of average level with threshold, as described in Section 4.3. A typical curve is shown in Fig. 8(*b*). By differentiation, a new curve may be obtained showing the fractional time spent by the envelope above various thresholds. Fig. 9 shows on linear scales this cumulative probability distribution of the noise, deduced from the curves of Fig. 8, and the corresponding distribution of fluctuation noise with the same mean level. The steep initial decrease of the curve for atmospheric noise and its subsequent long extension to high thresholds indicate a structure containing large impulses with relatively quiet intervals between them. The subsidiary graph shows the form of the distribution at very high thresholds, those values above thresholds of about 20 mV/m being deduced from the distribution of peaks.

It is known that the distribution of the envelope of fluctuation noise conforms to a Rayleigh law, given by

$$Q(v) = \exp(-v^2/2\bar{v}_0^2) \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where $Q(v)$ = Fraction of time spent above threshold v.
$2\bar{v}_0^2$ = Mean square voltage.

The atmospheric noise distribution has a different form, and attempts have been made to find a simple empirical law to represent it.

The large range of amplitudes present in the distribution suggests that they should be expressed in logarithmic units. Workers at the National Bureau of Standards and the University of Florida have found that the amplitude distribution can be represented reasonably well by a log-normal function,[8] i.e. the amplitudes expressed in logarithmic units are distributed normally. This function, with two independent parameters (the median and the standard deviation, for example) is more flexible

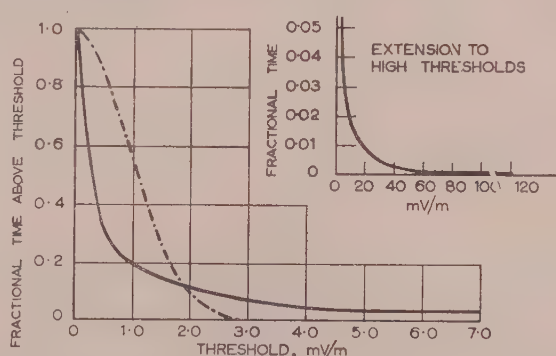


Fig. 9.—Cumulative amplitude probability distribution of atmospheric noise; 300 c/s bandwidth at 10 kc/s, 14th July, 1955.
——— Atmospheric noise.
—·—· Fluctuation noise of same average level.

than the Rayleigh function, which has only one, and the possibility of obtaining a good representation of the experimental distribution is thereby greatly increased. The curves shown in Fig. 9 have been replotted in Fig. 10(*a*) with the field strengths on a logarithmic scale and the probabilities on a 'normal distribution' scale. Most of the plots obtained in this way have been found to have slight curvature, but the departure from linearity may not have great practical significance. The data therefore indicate that the probability density may be represented by a function[8]

$$P(E) = \frac{1}{\sqrt{(2\pi)}\sigma E} \exp\left[-\frac{(\log E - \log E_m)^2}{2\sigma^2}\right] \quad . \quad (8)$$

where E_m (the median value of E) and σ (the standard deviation of $\log E$) are the independent parameters which define the distribution.

No theoretical justification for a log-normal distribution has been advanced, and the fact that the plots are linear does not provide an indication of the form of the noise. The distribution for fluctuation noise, also shown in Fig. 10(*a*), has only slight curvature, and it therefore appears that by introducing two

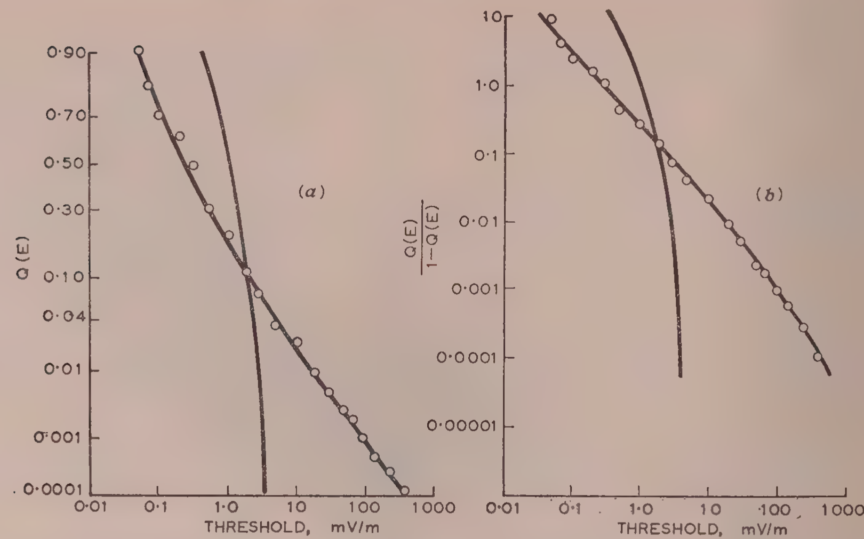


Fig. 10.—Cumulative amplitude probability distribution of atmospheric noise; 10 kc/s, 14th July, 1955.

—O—O—O— Atmospheric noise.
———— Fluctuation noise with same average field strength.

(*a*) Log-normal representation.
(*b*) Log-log representation.

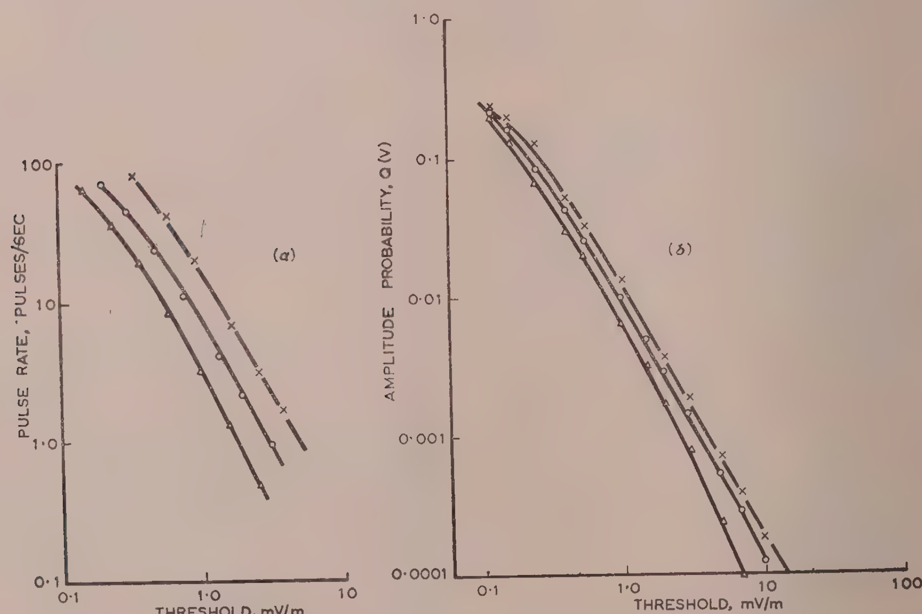


Fig. 11.—Variation of noise distribution with bandwidth; 10 kc/s, 1300–1500 G.M.T., 26th September, 1955.

—△—△— 190 c/s bandwidth.
—O—O— 310 c/s bandwidth.
—×—×— 540 c/s bandwidth.

(*a*) Pulse rate.
(*b*) Amplitude probability distribution.

variable parameters a wide range of types of noise can be represented by a log-normal distribution.

In view of the apparent lack of theoretical basis, the log-normal distribution should perhaps be treated with some caution. Other distributions with two variable parameters may provide as good a representation and with some theoretical justification. As an example, the data have been replotted in a different form in Fig. 10(*b*). Here the plotted distribution function is the ratio of the time for which the field strength lies above a given threshold to that spent below it, or $Q(E)/[1 - Q(E)]$ where $Q(E)$ is the cumulative probability distribution. From the approximate power-law relationship for the pulse rate [Fig. 9(*a*)] it can be shown that the plot in Fig. 10(*b*) will also be nearly linear at the high field strength levels, and it is in fact seen to be approxi-

mately linear over a large range of amplitude. The distribution can therefore be represented by the formula

$$Q(E) = \frac{C}{E^p + C} \quad . \; . \; . \; . \; . \; . \quad (9)$$

where C and p are constants.

In Section 6.4.1 the suggestion was made that the power-law relationship might prove to have some theoretical derivation based on storm distributions, and this would give some support to eqn. (9). On the other hand, the equation is a difficult one to manipulate mathematically, and must be terminated in a finite upper limit. Its validity as an empirical formula, however, does indicate that with two independent parameters other representations of the noise than the log-normal distribution are possible. Further work may lead to other formulae which are acceptable on both experimental and theoretical grounds.

(6.4.3) **Dependence of Structure on Bandwidth.**

In a previous Section the variation of average voltage with bandwidth was discussed. The series of measurements at various bandwidths also produced information about the variation of the structure. Fig. 11(*a*) shows how the pulse rates varied with bandwidth; the data are consistent with the assumption that the heights of individual peaks were proportional to the bandwidth. Fig. 11(*b*) shows the variation of the amplitude probability distribution Q(E) as a function of bandwidth. The distribution is seen to be similar over the range considered, the main change being a displacement along the amplitude axis. There is a slightly more than proportionate decrease in probability at the higher levels, and little change at the lowest levels plotted, the consequent steepening of the curve showing the tendency towards fluctuation noise. Even at 190 c/s bandwidth, however, large impulses are a prominent feature of the noise.

(6.5) The Time Distribution of the Major Impulses

It has been seen that large impulses form an important part of atmospheric noise at low frequencies. It is of interest to determine whether these impulses occur in random time-sequence, or whether there is some regularity or grouping in the occurrence. The time distribution was found from two types of sample; in one, the numbers of large impulses (exceeding a fixed high threshold) were counted in samples of about one minute, and in the other, impulses above various thresholds were counted in samples of about 10 millisec. Comparison of these time distributions with the Poisson distributions appropriate to random occurrences showed that, in the case of the longer samples and very large impulses, the occurrence was random. A study of the fine structure by smaller duration sampling from a film record indicated a departure from randomness, with a tendency towards occurrence in groups which might be expected from the knowledge that a lightning flash often consists of several strokes down the same channel. It is doubtful whether this tendency towards grouping would have much influence on the interfering properties of the noise to most types of service.

(7) CONCLUSIONS

Atmospheric noise at very low frequencies has been seen to consist of a succession of impulses, discrete at the high field-strength levels but interfering at the low levels to form something more resembling fluctuation noise. For most practical purposes the impulses can be regarded as occurring randomly in time.

The amplitude characteristics of the noise envelope may be expressed in terms of its amplitude probability distribution, but at the higher levels a rather more complete description may be given in terms of the amplitudes of the peaks, which have a form determined mainly by the bandwidth of the receiver. Either amplitude distribution can be described empirically by two parameters, the actual choice of the parameters being somewhat arbitrary. Both a log-normal law and a form of power law can describe the amplitude probability distribution satisfactorily, but theoretical justification for either of these representations is lacking.

As the bandwidth is varied the high-voltage regions of the noise behave as separate impulses; the peak heights are proportional to bandwidth for bandwidths down to 190 c/s and probably much lower. The average field strength, however, which is conditioned largely by the interference between the impulses at the low levels, is proportional to the square root of the bandwidth, for bandwidths up to 300 c/s. In some of its characteristics, therefore, the noise resembles a series of discrete impulses, while in others it is more like fluctuation noise. This emphasizes the necessity of using more than one parameter to describe the noise.

Analysis of the large number of observations which have been obtained by the technique described may assist in the choice of the best empirical representation of the noise, but a more detailed knowledge of the locations of thunderstorms, their power radiating properties and the propagation of very-low-frequency waves is required to give support to the results of noise measurements.

(8) ACKNOWLEDGMENTS

The work described above was carried out as part of the programme of the Radio Research Board. This paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

(9) REFERENCES

(1) Thomas, H. A., and Burgess, R. E.: 'Survey of Existing Information and Data on Radio Noise over the Frequency Range 1–30 Mc/s', Radio Research Special Report, No. 15, 1947 (H.M. Stationery Office), Code No. 47–29–15.

(2) Hogg, D.: 'The Measurement of Atmospheric Radio Noise in South Africa in the Low-Frequency Band', *Transactions of the South African Institution of Electrical Engineers*, 1950, **41**, p. 209.

(3) Gerson, N. C.: 'Noise Levels in the American Sub-Arctic', *Proceedings of the Institute of Radio Engineers*, 1950, **38**, p. 905.

(4) Carbenay, F.: 'Sur l'enregistrement du niveau moyen des atmosphériques', *Comptes Rendus*, 1950, **230**, p. 1648.

(5) Gardner, F. F.: 'The Use of Atmospherics to Study the Propagation of Very Long Radio Waves', *Philosophical Magazine*, 1950, **41**, p. 1259.

(6) Sullivan, A. W., van Valkenburg, H. M., and Barney, J. M.: 'Low-Frequency Atmospheric Radio Noise in Florida', *Transactions of the American Geophysical Union*, 1952, **33**, p. 650.

(7) Hoff, R. S., and Johnson, R. C.: 'A Statistical Approach to the Measurement of Atmospheric Noise', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 185.

(8) Foldes, G.: 'La XIe Assemblée Générale de l'Union Radio-Scientifique Internationale—Mesure de bruit atmosphérique (Commission IV)', *L'Onde Électrique*, 1955, **35**, p. 598.

621.396.61

Paper No. 2171 R
Nov. 1956

AN EXPERIMENTAL ASSESSMENT OF THE LINEARITY OF A V.H.F. TRANSMITTER

By D. E. HAMPTON, B.Sc.

(*The paper was first received 7th March, and in revised form 18th May, 1956.*)

SUMMARY

An experimental procedure is described for testing the assumption that a generator behaves as a linear source.

The source admittance is obtained from measurements made when the generator is operating normally, and the problem considered is that of matching this source admittance to the characteristic admittance of the feeder connecting it to a wide-band aerial.

Measurements were made on a transmitter designed to operate in the band 100–125 Mc/s, and from these results it is deduced that, with a given aerial and feeder, a considerable increase in output power may be obtained by modifying the source admittance of the transmitter.

LIST OF SYMBOLS

α, β = Constants of proportionality defined in Fig. 5.
B = Susceptance.
b = Normalized susceptance.
d = Distance between points $-Y_{L1}$ and $-Y_{L2}$ (see Fig. 5).
G = Conductance.
g = Normalized conductance.
I = Current from constant-current generator (r.m.s.).
ρ = Modulus of the voltage reflection coefficient between feeder and aerial.
s = Standing-wave ratio.
V = Voltage (r.m.s.).
Y = Admittance.
L, L_1, L_2, L_3 = Suffixes pertaining to loads.
S = Suffix pertaining to source.

(1) INTRODUCTION

The solution of electronic circuit problems involving non-linear elements usually presents difficulties, and generally an attempt is made to reduce such problems to ones involving linear elements alone. The particular problem of this type considered here is one in which a valve circuit acts as a generator (e.g. valve oscillators, resonant class-C amplifiers), and it is required to determine the behaviour of the generator for various loading conditions.

The problem generally arises in transmitters, where the power from the output stage of the transmitter unit must be fed to an aerial via a feeder line, and care must be taken in the design to avoid any serious mismatching of impedances. When the transmitter is designed to operate over a wide band of frequencies, it may be necessary to accept some degree of mismatch, in which case it is necessary to know how this affects the output power. If such a transmitter could be represented as a linear generator, a solution of the problem would be relatively simple, for once the source admittance and the short-circuit current were obtained the power generated in various loads could be calculated.

The paper describes an experimental method of obtaining these parameters of the equivalent linear generator from a few measurements made under normal operating conditions of load. The procedure consists essentially in measuring the behaviour of the transmitter under three separate loadings, and from these measurements obtaining values for the parameters. By repeating the experiment for various loads, new sets of parameters can be obtained and a comparison can be made to show how well the transmitter approximates to a linear generator over the range of loads used.

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
Mr. Hampton is at the Royal Aircraft Establishment.

(2) GENERAL PRINCIPLES

With the v.h.f. transmitter used in the experiments it was found that over the range of loads used there was no significant variation in frequency comparable with the variations usually found in microwave transmitters. Furthermore, from some preliminary measurements made of the powers delivered into various loads there appeared a strong similarity to the results expected from a linear source. It was this evidence that suggested that the transmitter be replaced by an equivalent linear generator.

The type of linear generator to be considered is a constant-current source shunted by an admittance, although a constant-voltage source in series with an impedance is completely equivalent by the duality of Thévenin's and Norton's theorems.[1] The primary object is to test the validity of replacing the transmitter by such a linear generator. There are three parameters needed to determine a linear generator, namely the short-circuit current and the real and imaginary parts of the source admittance. If the transmitter behaves as a linear generator, three equations relating these parameters to measurements made on three separate loadings will be sufficient to determine them. When the transmitter is not a perfectly linear source, different values for the parameters will be obtained from each set of three loading conditions, and it is the variation in these values which determines the accuracy to which the transmitter behaves as a linear source over the range of loadings used.

In practice it is convenient to measure the load admittance and the r.m.s. voltage developed across it. These are related to the r.m.s. short-circuit current and the source admittance of a linear generator by the equation

$$V_L = \frac{I}{|Y_S + Y_L|} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The simultaneous solution of three equations of this type for Y_S and I may be obtained analytically, but they are more easily obtained graphically by plotting the loads on a rectangular admittance diagram and carrying out certain geometrical constructions (see Section 8). When nine different loads are used and the results are divided into three sets of three, values for the source admittance may be obtained from each set. If these values are reasonably consistent, the transmitter may be assumed to behave as a linear generator over the range of loads used.

The linear generator derived will be equivalent to the actual source only in so far as the external behaviour is concerned. For example, there is no justification in comparing the internal losses of the linear generator with those of the transmitter: the internal behaviour of the transmitter could be obtained only by internal measurements.

The behaviour of a linear generator under various loadings is a relatively simple problem, although the author is unable to

ind a detailed account in the literature. It is well known that he maximum power is obtained when the load is the complex onjugate of the source admittance, and it is necessary in the roof of this to derive an equation for the power in an arbitrary oad.[1] This equation is of the form

$$P = \frac{G_L I^2}{(G_L + G_S)^2 + (B_L + B_S)^2} \quad . \quad . \quad . \quad (2)$$

vhere $G_L + jB_L$ is the load admittance, $G_S + jB_S$ is the source dmittance and I is the r.m.s. short-circuit current.

It is evident from this equation that the contours joining those oads which draw the same power from the transmitter form a amily of coaxial circles when the load admittances are plotted on a rectangular co-ordinate system. It is current practice in transmitter design to plot these contours on the polar form of admittance chart (Smith's chart) to give the Rieke diagram.[2] In this paper it has been found more convenient to use the rectangular form, because these circles are more easily constructed by calculating their radii and centres from eqn. (2), than by plotting three points on the Smith's chart and constructing the circle through them.

Typical examples of the appearance of the constant-power circles on a rectangular axis are shown in Fig. 1. The centres of these circles all lie on a line of constant susceptance, which is the negative of the source susceptance of the generator. The limit point of the system is the complex conjugate of the source

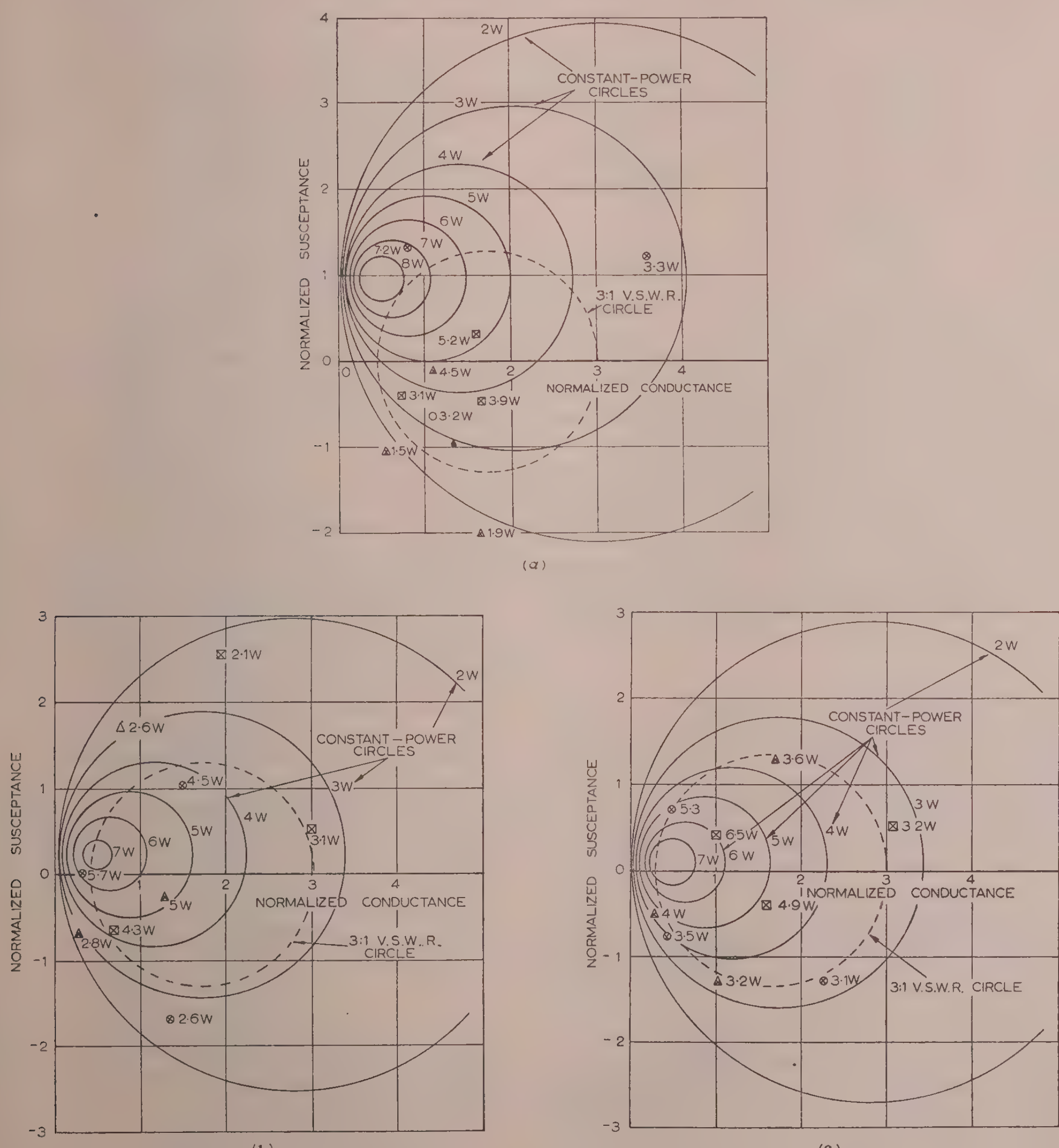


Fig. 1.—Output power of transmitter for various loads and frequencies.
(*a*) 101·9 Mc/s. (*b*) 116·1 Mc/s. *c*) 123·92 Mc/s.

admittance and is, of course, the load which dissipates the maximum power from the generator. As is customary, the loads have been normalized to the characteristic admittance of the feeder, assumed to be purely conductive, so that the constant-standing-wave-ratio coaxial circles have their centres on the real axis with unity as the limit point.

It has been pointed out before[2] that it is necessary to avoid the confusion that may arise in the terms 'matching of admittances' and 'matching of transmitter'. In order to match the transmitter, i.e. cause the transmitter to give its maximum power output, the aerial load presented to the feeder must be the complex conjugate of the source admittance. Unfortunately, owing to the practical situation of having different lengths of feeder for different installations, it is necessary to match the transmitter to the characteristic admittance of the feeder. This implies the introduction of a 4-terminal network between the transmitter and the feeder so that the source admittance of the transmitter appears as the characteristic admittance of the feeder. Alternately, this matching may be looked upon as making a load equal to the characteristic admittance of the feeder appear to the transmitter as the complex conjugate of its source admittance. When the transmitter is matched to the feeder the constant-power circles coincide with the constant v.s.w.r. circles and the power reaching the aerial is uniquely determined by the v.s.w.r. on the feeder by eqn. (3).

$$P = (1 - \rho^2)P_{max} = \frac{4S}{(S+1)^2}P_{max} \quad . \quad . \quad . \quad (3)$$

where P is the power reaching the aerial when the modulus of the reflection coefficient is ρ or the v.s.w.r. on the feeder is S. P_{max} is the power fed into a conjugate matched load.[4] This equation shows that when the transmitter is matched to the feeder there is only the mismatch between the aerial and the feeder to cause reflection of power. The incident power is the maximum output of the transmitter, P_{max}, while the reflected power is a proportion, ρ^2, of this. The total power dissipated in the aerial is the difference between the two. This is the formula usually given for calculating the loss in power due to a mismatched aerial, but it is not always explained that this presupposes that the transmitter is matched to the feeder.

(3) THE METHOD OF MEASUREMENT

The apparatus used for making the measurements consisted of (see Fig. 2) the transmitter under test and the load. The term 'load' refers to all the equipment across the output terminals of the transmitter and may itself be divided into the measuring section and the variable section. The former contained a slotted-line standing-wave indicator with a tunable probe unit and detector calibrated so that meter readings could be converte to values proportional to the voltage across the line at th position of the probe. The variable section varied the loa across the transmitter, and comprised an extending line ter minated by a wattmeter with a variable length of short-circuite line across it. The wattmeter formed that part of the load i which most of the power was dissipated and served as a con venient means of measuring this power.

From the values of the maximum and minimum voltage (V_{max}, V_{min}) and their positions on the line the load admittanc Y_L was obtained by the standard method.[2] The difference i

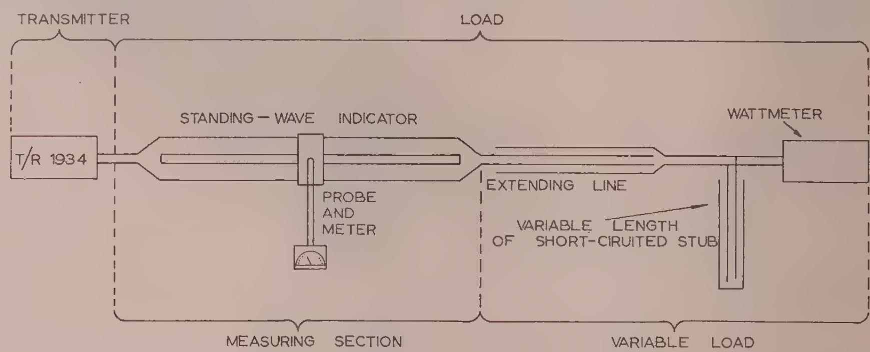


Fig. 2.—Experimental arrangement for the measurement of the source admittance of the transmitter.

position gave a check on frequency, and from the fact that th product of the maximum and minimum voltages is proportion to the power dissipated in the load,[3] the r.m.s. voltage acros the load was obtained as V_L, where

$$V_L \propto \frac{(V_{max}V_{min})^{1/2}}{\mathscr{R}(Y_L)} \quad . \quad . \quad . \quad . \quad . \quad (4$$

For the calculation of the source admittance it is sufficient t have values proportional to V_L and not necessarily their absolut values. When the source admittance had been obtained, value for the short-circuit current were deduced from the wattmete readings. By this method high accuracy could be achieved i the determination of the source admittance without relying o the wattmeter, which had an accuracy of about $\pm 10\%$.

(4) EXPERIMENTAL RESULTS

A transmitter with a push-pull class-C amplifier output stag as shown in Fig. 3 was used in the experiments, and measur ments and corresponding calculations of the source admittan were made at three frequencies. In making these calcul tions it was convenient to normalize the load admittan to the characteristic admittance of the slotted line, name

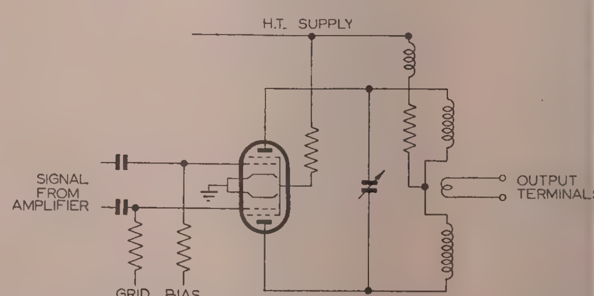


Fig. 3.—Output stage of transmitter.

9·2 millimhos (impedance of 52 ohms), so that in the results all dmittances must be multiplied by 19·2 to obtain their absolute alues.

Nine measurements made at 101·9 Mc/s were grouped into ree sets of three, and from each set a value of the source dmittance was obtained. These results are shown in the first art of Table 1, together with a marking characterizing the

Table 1

Frequency	Set marking	Calculated normalized source admittance
Mc/s		
101·9	Circles	0·45–j1·0
	Squares	0·43–j0·96
	Triangles	0·46–j1·02
116·1	Circles	0·46–j0·25
	Squares	0·47–j0·3
	Triangles	0·44–j0·23
123·92	Circles	0·42–j0·06
	Squares	0·46–j0·10
	Triangles	0·45–j0·12

t from which the calculation was made; the nine load dmittances are shown in Fig. 1(*a*), each point being circumribed by its appropriate set marking. Beside each point is e corresponding wattmeter reading, from which the shortrcuit current was calculated by taking a mean value determined om all the points. The constant-power circles shown drawn ere calculated from eqn. (2) for a normalized source admittance f 0·44–j0·98 and a short-circuit current of 540 mA. Also own is a dotted circle corresponding to a 3 : 1 v.s.w.r. for the otted line. Similar results are given for 116·1 and 123·92 Mc/s Table 1 and the corresponding Figs. 1(*b*) and 1(*c*). It can be en that the transmitter behaves as a linear source with a onductance which is practically constant but a susceptance which ecreases with increasing frequency.

From these results it is possible to analyse the variations in ower reaching an aerial under various degrees of mismatch. or example, suppose that at 101·9 Mc/s the aerial admittance is ch as to give a 3 : 1 v.s.w.r. on a 52-ohm feeder. The output the transmitter will depend on the admittance of the aerial resented to the length of feeder cable. This is the load Y_L presented to the transmitter, and it can be anywhere on the : 1 v.s.w.r. circle depending on the aerial admittance and the ngth of feeder. It is evident from Fig. 1(*a*) that the output of e transmitter varies considerably as the load Y_L moves round is circle. In Fig. 4 this variation is shown plotted against a ariation in the length of feeder. As the feeder is increased by /2 when supporting a 3 : 1 v.s.w.r. the power from the transitter varies between about 2 watts and just over 8 watts. Since e feeder length depends on the particular installation, it is ecessary to allow for the possibility that the power may be as w as 2 watts. Hence the minimum output power for a given anding-wave ratio should be specified as characteristic for ny transmitter and aerial combination. Table 2 gives such inimum values for the three frequencies considered here, for e transmitter used in the experiments.

If eqn. (3) is applied to the case when the transmitter is tuned 101·9 Mc/s, for which P_{max} is about 9 watts, the power at could be fed to the aerial giving a 3 : 1 mismatch is about 7 watts, from which it is evident that a considerable increase power could be achieved if some form of matching were troduced between the feeder and transmitter. An example the improvement possible in practice is also given in Table 2,

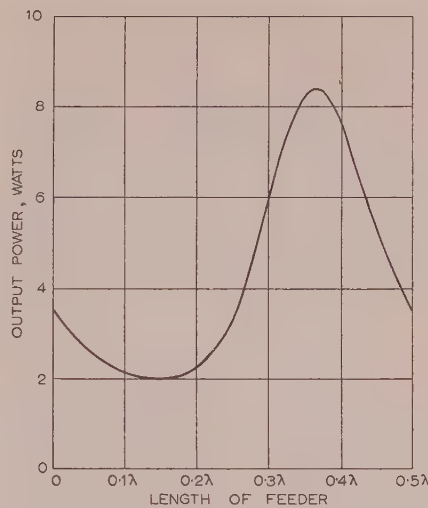


Fig. 4.—Variation of output power of transmitter with length of feeder.

3 : 1 v.s.w.r. on feeder.
Frequency = 101·9 Mc/s.

Table 2

MINIMUM CHARACTERISTIC OUTPUT POWERS OF TRANSMITTER-AERIAL COMBINATIONS

Frequency	Output power at various standing-wave ratios				
	1·5	2	3	4	5
Mc/s	watts	watts	watts	watts	watts
Experimental combination					
101·9	3·9	2·8	2·0	1·5	1·0
116·1	4·8	4·0	3·1	2·5	2·0
123·92	6·2	4·2	3·2	2·7	2·1
Improved combination					
101·9	6·5	5·5	4·3	3·4	2·9
116·1	7·0	6·2	5·0	4·0	3·6
123·92	6·5	5·8	4·6	3·8	3·3

a practical 4-terminal passive network having been introduced between the transmitter and the feeder to improve the matching over the band 100–125 Mc/s. These figures have been calculated to show what is possible, and it can be seen that powers previously obtained for a 1·5 : 1 v.s.w.r. could be achieved with a 3 : 1 v.s.w.r. under better matched transmitter conditions using a non-tunable matching unit.

(5) CONCLUSIONS

It has been shown experimentally that an essentially nonlinear source used in the experiment can be treated as a linear generator, and that the source admittance can be determined under normal operating conditions. By replacing the source by a linear generator, an analysis of the power delivered into any load is readily obtained. This simplifies the study of matching transmitters to wide-band aerials to obtain maximum radiated power.

(6) ACKNOWLEDGMENT

Acknowledgment is made to the Controller, Her Majesty's Stationery Office, for permission to publish the paper.

(7) REFERENCES

(1) SELGIN, P. J.: 'Electrical Transmission in Steady State' (McGraw Hill, New York, 1946), p. 37.
(2) BARLOW, H. E. M., and CULLEN, A. L.: 'Microwave Measurements' (Constable, London, 1950).
(3) WILLIS JACKSON: 'High-Frequency Transmission Lines' (Methuen, London, 1945).
(4) HUXLEY, L. G. H.: 'A Survey of the Principles and Practice of Wave Guides' (University Press, Cambridge, 1947), Chapter 5.
(5) PHILLIPS, E. G.: 'Functions of a Complex Variable' (Oliver and Boyd, Edinburgh, 1945).

(8) APPENDIX

A Graphical Solution for the Source Admittance

When a linear generator with a source admittance Y_S and r.m.s. short-circuit current I is terminated by a load admittance Y_{Li}, the r.m.s. voltage, V_{Li}, across the load is given by

$$V_{Li} = \frac{I}{|Y_S + Y_{Li}|} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

From three equations of the form (5) the current I may be eliminated to give

$$\frac{V_{L1}}{V_{L2}} = \frac{|Y_S + Y_{L2}|}{|Y_S + Y_{L1}|} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

$$\frac{V_{L2}}{V_{L3}} = \frac{|Y_S + Y_{L3}|}{|Y_S + Y_{L2}|} \quad . \quad . \quad . \quad . \quad . \quad (7)$$

The solutions for Y_S which satisfy eqn. (6) lie on a circle of the coaxial system which has $-Y_{L1}$, $-Y_{L2}$ as limit points.[5] A similar circle is defined by eqn. (7), and these two circles intersect in the two points $G_1 + jB_1$ and $G_2 + jB_2$. By solving the equations analytically it is possible to show that $G_1 + G_2 < 0$, and hence only one of the solutions can have a positive real part. This is the solution which represents the source admittance of the generator. These circles are constructed by making use of the properties of coaxial circles and their limit points. First, the centre lies on the line joining the limit points $-Y_{L1}$ and $-Y_{L2}$ and Fig. 5 shows the circle defined by eqn. (6). Let

$$\left.\begin{aligned} r_1 &= |Y_S + Y_{L1}| \\ r_2 &= |Y_S + Y_{L2}| \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (8)$$

If Y_S is a point on the circle, its distance from the limit points is given by r_1 and r_2 and the ratio of these distances, from eqn. (6), is such that

$$r_2 = \frac{V_{L1}}{V_{L2}} r_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

In particular, take the point where the circle meets the line joining $-Y_{L1}$ and $-Y_{L2}$, and let the distance between these

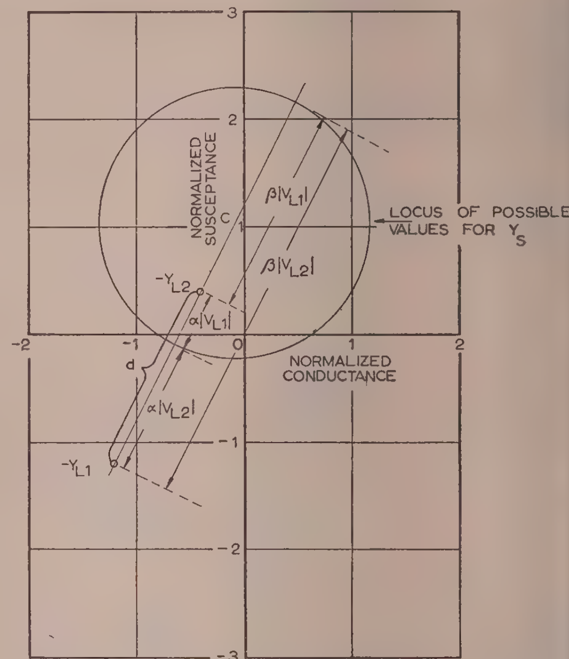


Fig. 5.—Graphical determination of source admittance.

$d = |Y_{L1} - Y_{L2}| \qquad \alpha = \frac{d}{|V_{L1}| + |V_{L2}|} \qquad \beta = \frac{d}{|V_{L2}| - |V_{L1}|}$

limit points (i.e. $|Y_{L1} - Y_{L2}|$) be d, as shown. Then for one point on the diameter,

$$r_1 + r_2 = d \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Solving eqns. (9) and (10) simultaneously gives

$$\left.\begin{aligned} r_1 &= \frac{dV_{L2}}{V_{L1} + V_{L2}} \\ r_2 &= \frac{dV_{L1}}{V_{L1} + V_{L2}} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (11)$$

Similarly, for the other point on the diameter,

$$r_1 - r_2 = d \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

which, combined with eqn. (9) gives,

$$\left.\begin{aligned} r_1 &= \frac{dV_{L2}}{V_{L1} - V_{L2}} \\ r_2 &= \frac{dV_{L1}}{V_{L1} - V_{L2}} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (13)$$

With these points determined the circle may be drawn and its intersection with a similar circle defined by eqn. (7) gives the required source admittance.

521.396.8.029.6

Paper No. 2184 R
Nov. 1956

POLARIZATION DISCRIMINATION IN V.H.F. RECEPTION

By J. A. SAXTON, D.Sc., Ph.D., Member, and B. N. HARDEN, M.Sc.

(*The paper was first received* 11*th May, and in revised form* 12*th July*, 1956.)

SUMMARY

An account is given of measurements in the band 40–200 Mc/s of he discrimination likely to be achievable between common-frequency ransmissions by the use of orthogonal polarizations. It is shown that he discrimination is determined primarily by the topographical nature f the receiving site, that it is substantially independent of distance rom the transmitter and of frequency in the band under consideration, nd that the median value is about 18 dB. The perturbing effects of ick-up on the feeder and of receiving aerial misalignment are discussed.

(1) INTRODUCTION

The limited availability of channels in the v.h.f. band neces-itates the operation of two, and sometimes more, transmitters n the same frequency if adequate coverage of an area such as he United Kingdom is to be achieved for television and sound roadcasting in this band. The problem of mutual interference n the respective service areas of such transmitters then arises, nd it is necessary to know at what spacing the transmitters hould be placed if the interference is not to exceed some specified olerable limit. The factor of greatest importance in determining he minimum spacing is the variation of propagation charac-eristics with meteorological conditions. This matter has been xhaustively studied, both in this country and in the United tates of America, and an account of the results of British work, nd the conclusions reached concerning station spacing, was iven in a paper* published a few years ago. The paper dealt rimarily with the spacing of high-power stations operating with he same wave polarization (vertical), although mention was nade of the fact that an extra discrimination of about 10 dB etween wanted and unwanted signals could be obtained by using orizontal polarization for the second of the two common-requency channels. This average value of 10 dB for the polariza-on discrimination factor was based on the somewhat sparse nformation obtained from limited observations at only one fre-uency in the television band I. In view of the worthwhile eduction in the spacing of transmitters which can be achieved if here is a reliable and substantial amount of discrimination ealized by the use of orthogonal polarizations, it was decided carry out a more extensive investigation of the problem, and he paper describes the result of this work, which was carried out frequencies in the range 40–200 Mc/s.

An evaluation of the possibilities of polarization discrimination equires ideally the measurement of (*a*) the horizontally polarized omponent of the field at any given point from a vertically olarized transmitter, and (*b*) the vertically polarized component om a horizontally polarized transmitter. Unless the receiving quipment is perfect for a linearly polarized field (either vertically horizontally), the discrimination actually observed under ractical conditions of reception will not necessarily be identical ith the ratio of the orthogonally polarized components of the diation field at the receiving aerial, and it may well vary ppreciably with the type of receiving aerial and feeder.

* SAXTON, J. A.: 'Long-Distance Propagation in relation to Television in the United ngdom', *Proceedings I.E.E.*, Paper No. 1270 R, April, 1952 (**99**, Part IIIA, p. 294).

Written contributions on papers published without being read at meetings are ited for consideration with a view to publication.
The paper is an official communication from the Radio Research Station, Depart-ent of Scientific and Industrial Research.

The discrimination obtainable in practice can also be influenced by the use of the directional properties of various types of aerial. This, however, is a matter which cannot readily be treated on a systematic basis, and the present work has been confined to obtaining, so far as possible, a statistical picture of the cross-polarized components of the radiation field.

(2) EXPERIMENTAL PROCEDURE

If the transmitter radiates initially a pure plane polarized wave (this is not necessarily so in practice) the resultant polarization at a given receiving site will depend upon (*a*) the rotation, if any, of the plane of polarization during propagation through the troposphere, and (*b*) any changes due to the effect of the ground on propagation, including diffraction by and reradiation from obstacles such as trees, buildings and general irregularities in the terrain. With a perfect receiving installation, i.e. one in which the aerial (assumed to consist of a single linear element) picks up only vertically or horizontally polarized waves according to its orientation and in which no voltage is induced in any other part of the equipment (e.g. the feeder), there would be no uncer-tainty in the measurement of the vertically polarized component of the field. The same is not true, however, of the horizontally polarized component, for here the directional properties of the receiving aerial are of importance: it would theoretically be possible to obtain any value of signal between zero and a maxi-mum value as the aerial is rotated. In the present investigation a simple half-wave dipole was used for reception, and for the measurement of horizontally polarized fields it was always placed broadside-on to the direction of the transmitter. Although for vertically polarized transmissions this was often the condition for maximum horizontal signal, and therefore did not correspond to the greatest achievable discrimination, the procedure seemed the best one to adopt, since for the purpose of statistical analysis a standard arrangement was desirable; furthermore, it also covered the worst case of interference likely to arise in practice when the wanted and unwanted signals arrive from the same (or reciprocal) direction, and when it might not therefore always be possible to make use of the directional properties of the receiving aerial to increase the discrimination.

The conditions of reception normally encountered are, of course, not perfect, and the possibility of pick-up by the feeder must be taken into account; the influence of these factors on the observed polarization discrimination is discussed in Section 3.1.

The receiving dipole was connected by a balanced screened-twin feeder through a wide-band balance-to-unbalance trans-former to the unbalanced input terminals of a calibrated receiver. The dipole was mounted on a wooden mast so attached to a motor van that it could easily be raised to a height of 9 m above ground level, at which height all of the measurements were made. Facilities were also provided for rotating the aerial to any desired bearing, and for fixing it accurately in either the vertical or hori-zontal position. An overall calibration of the receiving equip-ment, carried out with a local transmitter on a good open site where it was reasonable to assume that no spurious effects would occur, showed that the signal picked up by the feeder alone was about 40 dB below that picked up by the aerial in the normal receiving position, both for horizontally and for vertically

polarized waves. It was with these experimental arrangements that most of the field observations were made, but some measurements were also taken at frequencies in bands II and III of the difference in discrimination obtained when an unbalanced coaxial feeder was substituted for the balanced screened-twin transmission line.

Observations were taken at a number of sites in open and built-up areas in southern England, using vertically polarized television sound transmissions from B.B.C. stations operating in band I, from Alexandra Palace at 41·5 Mc/s, from Sutton Coldfield at 58·25 Mc/s, and from Wenvoe at 63·25 Mc/s; and from the I.T.A. station, Croydon, at 191·25 Mc/s in band III. Horizontally polarized sound transmissions were obtained from the B.B.C. f.m. transmitters, Wrotham, at 89·1 and 93·5 Mc/s in band II.

The procedure at each site was to take three sets of observations with the dipole vertical and then horizontal for any particular frequency and to record signal variations in each case for periods of up to 3 min. From these records median values of field strength for each polarization and period were determined. For local transmitters the field strength did not vary appreciably with time, but with distant stations tropospheric effects could on occasion cause fading of up to 30 dB over short time intervals: it was considered, however, that this method of measurement would yield reasonably consistent values and that there was no need for more prolonged measurements. In addition, observation over a period of a few minutes helped to eliminate small variations due to movement of the aerial in the wind.

The sites and times of observation were not the same for all frequency bands, but measurements on different frequencies in each band were made together. A large number of sites were investigated in each case, so that comparison between results on frequencies in different bands is possible.

Since the measuring equipment was carried in a motor van, all of the observations were made with the receiving aerial over the road, but the height of the aerial was comparable with that normally used for domestic television reception. In urban areas the van was stationed at the side of the road with sufficient clearance from any nearby trees and overhead wires for the mast to be raised; the location of the aerial with respect to houses, trees and wires varied over a wide range of conditions.

(3) EXPERIMENTAL RESULTS AND DISCUSSION

Table 1 gives details of the transmissions investigated, the distances over which measurements were made, the number of

Table 1

Band	Frequency	Polarization	Number of sites	Distance	Discrimination exceeded at 10% of sites	50%	90%
	Mc/s			km	dB	dB	dB
I	41·5	Vertical	55	19–62	27	18	13
	58·25	Vertical	55	133–176	24	18	11
	63·25	Vertical	55	155–208	22	16	10
II	89·1	Horizontal	116	43–120	28	19	12
	93·5	Horizontal	116	43–120	28	19	12
III	191·25	Vertical	123	14–60	30	20	12

sites examined in each case, and also the values of the polarization discrimination factor exceeded at 10, 50 and 90% of the sites.

Fig. 1(*a*) shows the results for the vertically polarized band I transmissions; this set of measurements contains some made at

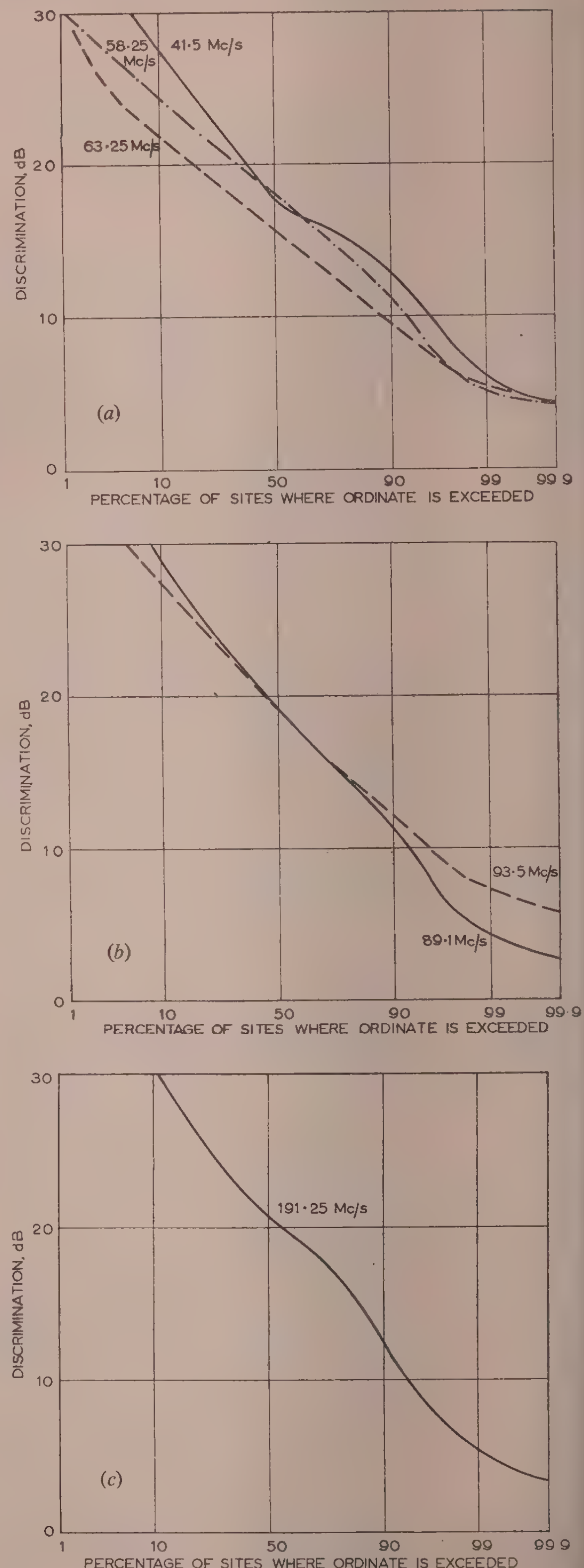


Fig. 1.—Polarization discrimination.
(*a*) Band I. (*b*) Band II. (*c*) Band III.

distances as great as 208 km—sufficiently great, in fact, for the propagation to be affected by variations in the refractive-index structure of the troposphere. However, it was found that the range of discrimination observed for each frequency at the more distant sites did not differ significantly from that observed at much shorter distances, and it may thus be concluded that any change of the plane of polarization due to propagation through the troposphere is negligible in comparison with changes due to other causes. There may be an indication of a small deterioration in discrimination at long distances, but variations in field strength arising from tropospheric effects decreased the accuracy with which the discrimination could be determined; a further disturbing influence was the presence of noise, which made it difficult to obtain a precise measurement of the weak cross-polarized component at such distances. It would therefore be unwise to attach any real importance to this rather slight tendency in the results.

Figs. 1(*b*) and 1(*c*) show the corresponding data obtained at band II and band III frequencies, and if these are examined together with Fig. 1(*a*) it will also be seen that there is no significant dependence of the discrimination factor on frequency over the band 40–200 Mc/s. It is further apparent that the discrimination factor is to all intents and purposes the same whether the transmission is initially vertically or horizontally polarized.

It would therefore appear that the discrimination factor depends mainly on the nature of the receiving site, and that variations in the factor are to be attributed to interaction between the field components due to diffraction by and reradiation from local obstacles and general irregularities in the terrain. High discrimination was obtained on carefully chosen rural sites, clear of trees and buildings, but a high value could also be obtained in urban areas where, presumably, the secondary components of the field due to obstacles combined to give a low value of the cross-polarized component. Low discrimination factors were observed on rural sites, particularly in the neighbourhood of trees, as well as in urban areas. Thus, in general, no marked distinction in overall discrimination characteristics was obtained between rural and urban conditions.

At any particular receiving site the discrimination varied widely with frequency and with small changes in the position of the receiving aerial. For example, the band II measurements were taken with a dipole tuned for the frequency midway between 89·1 and 93·5 Mc/s, and Fig. 2 shows how the difference in the

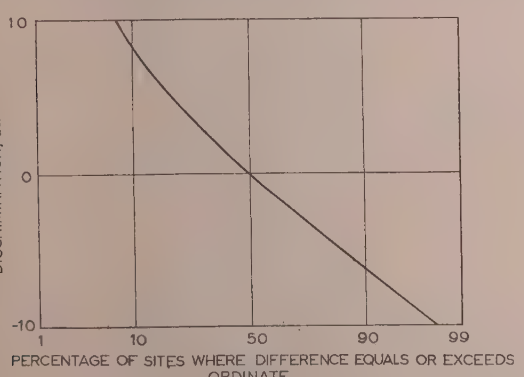


Fig. 2.—Difference in discrimination at two frequencies in band II.

discrimination at the two frequencies varied over all of the sites examined; although differences in discrimination of up to 16 dB were observed on some sites, the median values were the same. Similarly, in band I, large variations in discrimination were observed at the different frequencies on any given site. Changes in the position of the receiving aerial of only a few metres on some sites resulted in variations of as much as 15–20 dB in the discrimination.

(3.1) Effects of Pick-Up on the Feeder and Aerial Misalignment

Suppose that, in a field having vertically and horizontally polarized components, the input voltages to the receiver due to pick-up on the dipole alone are V_V and V_H respectively when the aerial is placed accurately vertical and horizontal. Suppose further that the input voltage due to pick-up on the feeder is V_F. The effective discrimination, D, will be different from V_V/V_H, and will depend upon the phase relationships between V_V, V_H and V_F. Consider first the case of a vertically polarized transmission: Fig. 3 then shows, for a series of values of V_F respec-

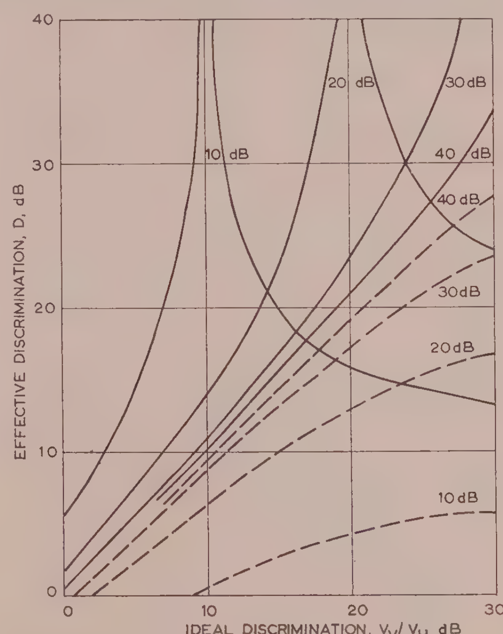


Fig. 3.—Effect of feeder pick-up on discrimination.

Effective discrimination as a function of V_V/V_H for ratios of V_V/V_F of 10, 20, 30 and 40 dB.

—— Maximum.
---- Minimum.

tively 10, 20, 30 and 40 dB below V_V, the curves of maximum and minimum values of D as a function of V_V/V_H, the discrimination which would be obtained with an ideal system. It will be seen that, provided V_F is at least 40 dB below V_V, the effective discrimination is very close to V_V/V_H; but if the pick-up on the feeder becomes relatively greater, D can differ considerably from V_V/V_H either way. The curves can be applied to initially horizontally polarized transmissions by interchanging the suffixes V and H.

A few measurements were made at band II and band III frequencies of the discrimination obtained at typical sites when an unbalanced coaxial feeder was substituted for the balanced screened-twin feeder. As mentioned previously, the pick-up on the latter feeder was known to be some 40 dB below that on the dipole when oriented for the appropriate wave polarization; the observations made with the twin feeder could therefore be regarded as a measure of V_V/V_H to a close approximation. For the horizontally polarized transmissions in band II the effective discriminations observed using the coaxial feeder at 13 sites were

scattered over a range of values lying between the maximum and minimum limits corresponding to $V_H/V_F = 25$ dB: at each site different discriminations were obtained at the two frequencies of 89·1 and 93·5 Mc/s. For the vertically polarized band III transmissions on 191·25 Mc/s the effective discriminations measured at 19 sites lay within the region defined by the maximum and minimum limits corresponding to $V_V/V_F \simeq 15$ dB. It should be noted that, as V_V/V_F decreases it is to be expected that increasingly more values of D would lie below V_V/V_H than above it, and this tendency was noticeable with the band III measurements.

The other factor having an influence on the effective discrimination which must be considered is that of aerial misalignment. If the dipole is not accurately vertical or horizontal, as the case may be, according to the polarization of the wanted signal, the observed discrimination will again differ from V_V/V_H (or V_H/V_V). The order of magnitude of this effect is illustrated by Fig. 4, in

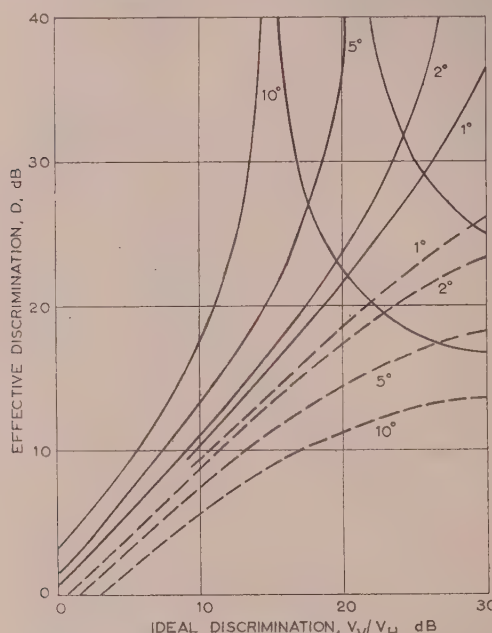


Fig. 4.—Effect of aerial misalignment on discrimination.

Effective discrimination as a function of V_V/V_H for angles of deviation of the dipole of 1, 2, 5 and 10° from the true position.
—— Maximum.
---- Minimum.

the compilation of which it has been assumed that the pick-up on the feeder is zero. The maximum and minimum values of D, depending upon the relative phases of V_V and V_H, which can occur for angles of deviation from the vertical and horizontal positions in either direction of 1, 2, 5 and 10° are shown as a function of V_V/V_H, the radiation here being supposed initially vertically polarized. The curves are also applicable to V_H/V_V if the transmission is horizontally polarized.

In a practical installation, of course, both feeder pick-up and aerial misalignment may be present, and both will make their contribution to the difference between V_V/V_H (or V_H/V_V) and the actual discrimination.

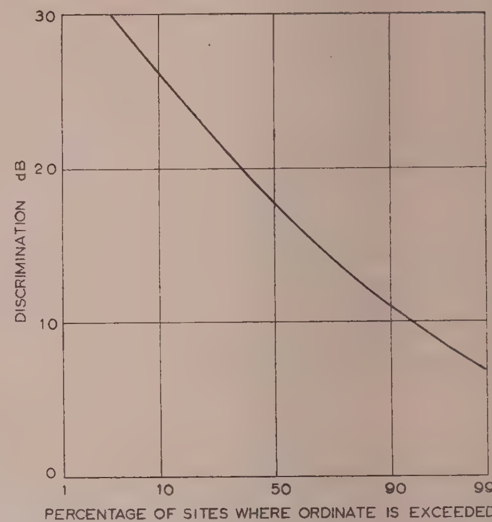


Fig. 5.—Basic polarization discrimination factor in the v.h.f. band.

(4) CONCLUSIONS

It appears that the discrimination achievable between common frequency transmissions by the use of orthogonal polarizations is largely independent of frequency in the band 40–200 Mc/s. The value of the polarization discrimination factor depends primarily on the perturbation of the field due to diffraction and reradiation by obstacles in the immediate neighbourhood of the receiving aerial, and does not vary significantly with distance from the transmitter. There is no evidence of rotation of the plane of polarization during propagation through the troposphere, and the discrimination appears to be much the same whether the wanted field is vertically polarized and the unwanted field horizontally polarized, or vice versa. Fig. 5 gives a mean curve of all the observations at various frequencies between 40 and 200 Mc/s, and it should provide a basic guide as to the order of the polarization discrimination factor which may be achieved in practice. It must be remembered, however, that the discrimination actually obtained in any particular instance will be influenced by such factors as pick-up on the feeder, aerial misalignment and the directional properties of the receiving aerial, the importance of which it is difficult to assess in any definite manner. They may produce effective discriminations either greater or less than those given in Fig. 5, but it should be noted that, when the pick-up on the feeder is a significant fraction of that on the aerial, there will be a definite trend towards smaller discrimination.

(5) ACKNOWLEDGMENTS

The work described was carried out as part of the programme of the Radio Research Board. The paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

1.396.11

Paper No. 2186 R
Nov. 1956

SOME AIRCRAFT MEASUREMENTS OF BEYOND-THE-HORIZON PROPAGATION PHENOMENA AT 91·3 Mc/s

By B. J. STARKEY, Dipl.Ing., Associate Member.

(*The paper was first received* 1st *May, and in revised form* 17*th July*, 1956.)

SUMMARY

Field-strength measurements at distances extending far beyond the rizon from a transmitter on a frequency of 91·3 Mc/s have been rried out in an aircraft flying at a height of 10 000 ft.

The analysis of the results obtained and their correlation with teorological data suggest that many phenomena of long-distance opagation could possibly be explained by the simple hypothesis of ecular reflection from temperature-inversion layers at the tropopause.

(1) INTRODUCTION

In the course of some field-strength measurements carried out an aircraft at distances extending far beyond the horizon it s found that the results obtained cannot be readily explained terms of the usually suggested mechanisms of radio wave opagation. The paper gives a preliminary description of the als, which are still in progress, and attempts to interpret the sults.

(2) DESCRIPTION OF THE EXPERIMENT

Field-strength measurements were carried out in an aircraft ing at a height of 10 000 ft, from a transmitter on a frequency 91·3 Mc/s.

Special C.W. radiations were provided by some of the B.B.C. transmitters at Wrotham, the effective radiated power being of the order of 120 kW with horizontal polarization, and the receiving equipment was obtained on loan from the Post Office Engineering Department. The receiver output, read from a calibrated meter, was recorded at discrete intervals. Owing to the narrow-band low-pass filter characteristics of the receiving equipment, any fast fluctuations of the received signal were not recorded.

The following characteristics were found to be a common feature in the detailed structure of all the results of several flights. The results of one flight are plotted in Fig. 1 as an example.

(*a*) The initial rate of decrease of field strength beyond the horizon is, in general, slightly slower than that obtained from the simple diffraction theory for standard atmosphere.

(*b*) Two more or less pronounced maxima occur in the regions A and B (Fig. 1), at ranges of roughly 250 and 300 miles from the transmitter.

(*c*) A pattern of fairly deep and regular fades often appears beyond a range of about 350 miles (C, Fig. 1), in contrast to the relatively much shallower fluctuations of similar spacing occurring at shorter ranges.

(*d*) The signal drops rapidly by about 8 dB at 420 miles (D, Fig. 1) and fluctuates at about the noise level beyond that range.

Written contributions on papers published without being read at meetings are ited for consideration with a view to publication.

Mr. Starkey is at the Royal Aircraft Establishment, Ministry of Supply.

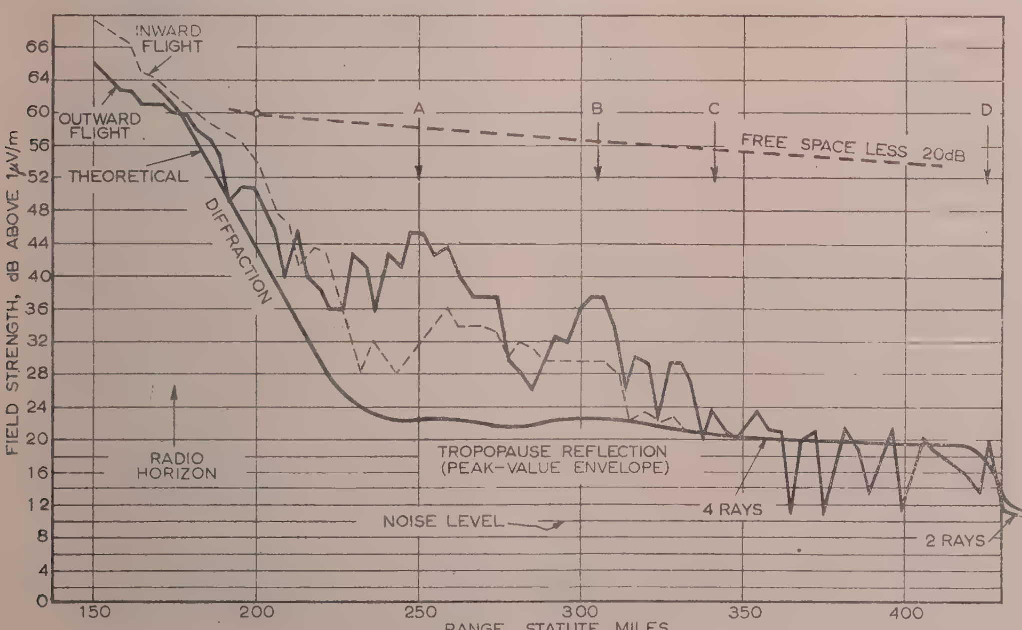


Fig. 1.—Variation of field strength with range.

(3) DISCUSSION OF THE RESULTS

The phenomena of persistent beyond-the-horizon propagation of v.h.f. and s.f. radiation are usually attributed either to 'forward scatter' by a turbulent atmosphere[1,2] or to partial reflections in an atmosphere possessing a gradient of refractive index.[3,4] The characteristics enumerated above, however, cannot be satisfactorily explained by either of the existing theories, although at least some of these characteristics seem to be present in all the intensity-range graphs published by other investigators.[1,5]

A study was therefore made to determine whether another mechanism of propagation could be found which better fitted the experimental results. One possible mechanism could be specular reflections from elevated layers.[6] It will be shown that a layer at a height of about 40 000 ft, with a sharp boundary between two media with refractive indices differing by about 2×10^{-6}, provides at least a partial answer to the problem. It so happens that the existence of such a layer does not have to be postulated, because a temperature inversion occurring in the tropopause is, in fact, a characteristic feature of the atmosphere in this country between 30 000 and 40 000 ft, and there are occasional secondary inversion layers at heights between about 20 000 and 50 000 ft.[7]

Fig. 2 shows the temperature/height relation, measured by one of the twelve meteorological stations which normally carry out

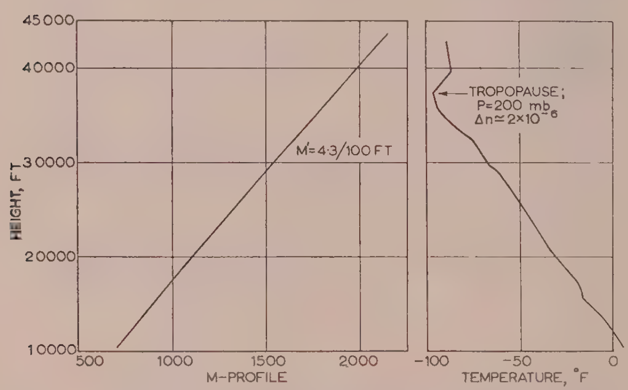


Fig. 2.—Variation of M-profile and temperature with height; Lerwick, 12th December, 1955, 1349 hours G.M.T.

such measurements in this country, on the same day as the flight corresponding to Fig. 1. The inversion is seen to occur at a height of about 38 000 ft, and the calculated refractive-index change at the inversion is here $\Delta n = 2 \times 10^{-6}$. Unfortunately, the available meteorological measurements are insufficiently accurate to permit determination of the thickness of the inversion layer, since the average height intervals in these measurements are of the order of 1 000 ft. There are indications, however, that the inversion is sharp, so that the reduction of the reflection coefficient, given for grazing incidence by

$$|R| \simeq \frac{\Delta n}{2\phi^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

(where ϕ is the angle which the incident wave makes with the boundary) may be negligible, at least at metre wavelengths.

In the presence of an elevated layer the radiated wave can be propagated to a receiver situated beyond the horizon via four paths, as shown in Fig. 3.

Fig. 4 serves to explain the geometry involved in the computation of maximum ranges achievable for individual rays; it can be seen that rays Nos. 1 and 2 suffer diffraction for receiver

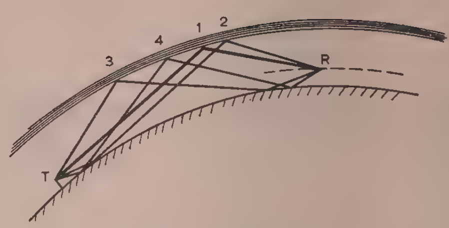


Fig. 3.—The four propagation paths involving one reflection from a elevated inversion layer.

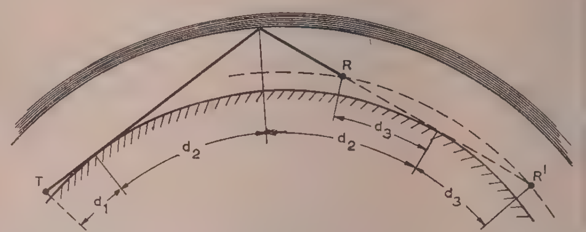


Fig. 4.—Maximum propagation range involving one reflection from a elevated inversion layer.

TR = Path for rays 1 and 2.
TR′ = Path for rays 3 and 4.
d_1 = 40 miles.
d_2 = 260 miles.
d_3 = 130 miles.

positions beyond R so that their contribution to the field strengt at points beyond R becomes negligible (apart from a limite region about R′ where ray 1 reappears). The remaining ray are present, however, up to R′ before they are diffracted.

From the simple geometry of Fig. 4 and for the termina heights involved, the distance T–R can be computed for a reflect ing layer at 40 000 ft and for a 'mean' effective earth radiu reduced for elevated paths by about 10% compared with that o a standard atmosphere, in order to take into account the actua refractive-index gradient at high altitudes. This distance is abou 420 miles, which is the range of D in Fig. 1. The maximur possible range for any rays, T–R′, is about 700 miles.

The field strength of individual rays can be computed from th geometry of Fig. 3, using the appropriate reflection coefficient a the reflecting layer and introducing the corrective convergenc factor for reflections from the concave elevated layer and th divergence factor for reflections from the convex earth. Th sum of the field strengths of four and two rays, respectively representing the peak value of the total signal, has been plotte in Fig. 1.

It can be seen that the experimental and theoretical results agre fairly well, at least quantitatively, for distances greater than abou 340 miles. At shorter distances the theoretical values are to low. It should be mentioned, however, that at these shorte distances (200–300 miles) experimental results vary greatly nc only from trial to trial, occasionally down to the level of th theoretical peak-value curve, but even within a period of a fe hours, represented by the time difference between the outwar and inward flights, as seen in Fig. 1. These variations might b due, perhaps, to the vagaries and irregularities at the tropopause on the other hand, they could be caused by contributions fror the lower atmosphere. An inversion layer at 6 000 ft was in fa present on the day in question. Only the equivalents of ray 3 and 4 in Fig. 3 could reach the aircraft flying above this laye and their effective range is about 340 miles, which correspond to C in Fig. 1.

This agreement could be taken as an indication that, at least n this particular occasion, the low-level inversion layer rather han any other mechanism was mainly responsible for the results medium ranges.

As stated before, however, the absence of sufficiently accurate eteorological data precludes any more direct approach to the roblem, which depends on the detailed structure of the lower mosphere.

The fluctuation patterns appearing in the experimental results ould be explained by interference between various rays. For ample, the deep regular fades beyond C can be due to interrence between the tropopause-reflected rays, with no other gnal blurring their appearance. It can be shown that in the gion C–D, i.e. close to the diffraction range of rays 1 and 2, e interference depends mostly on rays 1 and 3.

If the spacing between two adjacent maxima of such an interrence pattern is d and the terminal heights are h_1 and h_2, then e necessary height of the reflecting layer, h_i, can be calculated flat-earth approximation from

$$h_i \simeq \frac{\lambda D^2}{4dh_2} + \frac{h_1}{2}$$

or $D \simeq 400$ miles, $d = 8$ miles, $\lambda = 3{\cdot}3$ m, and the terminal eights involved, we get $h_i \simeq 30\,000$ ft. For spherical earth the sult would be greater. The order of magnitude is again constent with the height of the tropopause.

It can be said in conclusion that the simple hypothesis of ecular reflection from inversion layers at the tropopause resents a possible explanation of many phenomena of longistance propagation.

(4) ACKNOWLEDGMENTS

The author is indebted to the B.B.C. and to the Post Office ngineering Department—in particular to Mr. R. A. Rowden of the B.B.C. Research staff and Mr. J. K. S. Jowett of the Post Office—for their willing co-operation and assistance in the experiments. The flight measurements were carried out by Mr. G. K. Kitchen in an aircraft of the A. and I.E.U. Martlesham, and the analytical work by Mr. W. R. Turner, both of the R.A.E.

The whole project was undertaken and completed thanks to Dr. J. S. McPetrie, of the R.A.E., Chairman of the Tropospheric Wave Propagation Committee, D.S.I.R. Acknowledgment is due to the Controller, H.M. Stationery Office, for permission to publish the paper.

(5) REFERENCES

(1) MEGAW, E. C. S.: 'Scattering of Electromagnetic Waves by Atmospheric Turbulence', *Nature*, 1950, **166**, p. 1100.
(2) BOOKER, H. G.: 'A Theory of Radio Scattering in the Troposphere', *Proceedings of the Institute of Radio Engineers*, April, 1950, **38**, p. 401.
(3) FEINSTEIN, J.: 'Tropospheric Propagation beyond the Horizon', *Journal of Applied Physics*, 1951, **22**, p. 1292.
(4) CARROLL, T. J., and RING, R. M.: 'Propagation of Short Radio Waves in a Normal Stratified Troposphere', *Proceedings of the Institute of Radio Engineers*, 1955, **43**, p. 1384.
(5) AMES, L. A., NEWMAN, P., and ROGERS, T. F.: 'V.H.F. Tropospheric Over-Water Measurements far beyond the Radio Horizon', *ibid.*, p. 1369.
(6) NORTHOVER, F. H.: 'The Anomalous Propagation of Radio Waves in the 1–10 m Band', *Journal of Atmospheric Physics*, 1952, **2**, p. 106.
(7) BANNON, J. K., FRITH, B. A., and SHELLARD, H. C.: 'Humidity of the Upper Troposphere and Lower Stratosphere over Southern England', *Meteorological Office, Geophysical Memoirs* No. 88, London, 1952.
(8) SAWYER, T. S.: 'Day-to-Day Variations in the Tropopause', *ibid.*, No. 92, London, 1954.

DISCUSSION ON 'THE VERTICAL RADIATION PATTERNS OF MEDIUM-WAVE BROADCASTING AERIALS'*

Mr. K. L. Rao (*India; communicated*): This excellent paper as come at a time when only meagre data are available on the ffects of ground irregularities on the v.r.p. of a vertical radiator. he common belief that, so long as there are no 'shadows' cast y distant hills, any fairly even ground would do for the broadasting aerial does not appear to hold. The authors have condered at great length the effect of site irregularities on the v.r.p. f a single radiating mast, and it may be worth while to consider e effect on a two-mast or three-mast radiating system. It may e that directional systems are not in use in the United Kingdom, ut all the same the results for the directional systems may have nportant applications elsewhere.

What would be the effect on the radiating pattern of the vertical diator when situated in a valley ranged on either side by hills? Vould a Beverage type of aerial be the answer for such a site?

Messrs. H. Page and **G. D. Monteath** (*in reply*): We do not ink that ground irregularities *per se* will have much effect on the horizontal radiation patterns of directional aerial systems Hills comprising rock of low conductivity, or covered with forest, may however increase the attenuation of the ground wave. The extent to which energy is scattered at high angles by irregularities will, of course, depend on whether the irregularities are strongly illuminated by a directional aerial.

The vertical radiation pattern of a vertical radiator in a valley would probably be impaired rather less in the direction of the valley than in perpendicular directions, particularly if the cross-section of the valley was fairly uniform. The pattern could be calculated by an extension of the method given in the Appendix of our paper, provided the differences in height involved were small compared with the wavelength. A directional aerial, arranged to avoid radiation towards the hills, may well give the best performance in a long straight valley. Beverage aerials are inherently inefficient, depending on ground loss for their operation, and are normally used for reception only. Moreover, the vertical radiation pattern would, in general, be less suitable for a broadcasting service than that of a vertical aerial.

* PAGE, H., and MONTEATH, G. D.: Paper No. 1714 R, September, 1954 (see 102 B, 279).

621.396.677.45

Paper No. 2170 I
Nov. 195

AN EXPERIMENTAL DESIGN STUDY OF SOME S- AND X-BAND HELICAL AERIAL SYSTEMS

By G. C. JONES, Associate Member.

(*The paper was first received* 7*th March, and in revised form* 15*th June*, 1956.)

SUMMARY

The general characteristics of helical aerials are summarized and the results of polar-diagram measurements in the S- and X-bands are given for single and multiple helices. Reference is made to methods of widening both the beam width and the bandwidth obtained with a helical aerial by means such as end loading and tapering.

The possibility of adapting helical aerials to give wide-band linearly polarized arrays is explored experimentally with a considerable measure of success.

Tests show that acceptable characteristics may be obtained with helices of 'unorthodox' form or with dimensions outside the limits suggested by Kraus.

Encapsulation of helical aerials in foamed dielectric material in order to improve their rigidity is found to be a satisfactory and practical proposition provided that the materials are selected with some care, particularly in the X-band.

(1) INTRODUCTION

The characteristics of helical aerials for the radiation of circularly polarized beams have been studied at considerable length by a number of workers, and there is a fairly extensive literature on the subject. Practical information, however, is largely confined to frequencies below 400 Mc/s.

The work described here deals with tests carried out on helical aerials in the S- and X-bands, i.e. 10 cm and 3 cm, and also covers investigation of helices which may be considered as 'unorthodox' in that they depart from the normal ratios of circumference and pitch to wavelength, or take a rectangular instead of circular cross-section. The possibility of using helical aerials to give wide-band linearly polarized arrays in the S-band and the 300 Mc/s band has also been investigated.

This study is almost wholly confined to the aspect of the polar diagrams obtainable and does not bear on the impedance or matching of helical aerials to any extent, although it has been found[9] that constancy of impedance is closely related to the maintenance of good polar diagrams. Certain of the References quoted in the paper, and particularly Reference 9, deal with the impedance characteristics of helical aerials.

(2) GENERAL PROPERTIES

Circular polarization is commonly used where the requirement is for polarization diversity and a broad frequency band, or the avoidance of accurate alignment of receive and transmit aerials. Interference from unwanted reflections can also be reduced by the use of circular polarization as the direction of rotation of polarization is reversed at a reflecting plane, and a left-hand or a right-hand circularly polarized aerial will not respond to polarization of the opposite rotational sense.

Although other methods of obtaining circular polarization are available, such as by the use of circularizers with linear radiators, the adoption of helical aerials has several advantages, including greater bandwidth and non-critical construction.

Several papers dealing with helical aerials have been publishe by Kraus[1] and his co-workers in which extensive experiment results are presented and the results of radiation field-measure ments are correlated with an approximate theory. Kraus ha determined experimentally that the current distribution along helical wire may reasonably be approximated by a travellin sinusoidal wave whose phase velocity varies with frequency ov the useful range in almost precisely the manner necessary t provide the phase difference between turns required for an enc fire array of maximum directivity. It has been shown by Hanse and Woodyard[2] that maximum directivity in an end-fire arra is obtained with a phase difference between successive elemen slightly different from that required to cause the fields of eac element to add in phase along the line of array. It is remarkab and extremely fortunate that this condition should be so closel approximated over such a broad frequency band with a helic aerial. Kraus's approximate calculation of the radiation fiel was obtained by considering such a wave of current on a heli of square turns, the field of a single square turn being calculate first and then multiplied by an array factor appropriate to th number of turns in the helix, their spacing and their phas difference.

Kornhauser[3] has developed a rigorous formula for the radia tion field of a helical aerial derived on the assumption of th empirical current distribution obtained by Kraus. For a heli of several turns this formula yields results very nearly the sam as those obtained by Kraus, and it has the additional advantag of greater simplicity of computation and applicability to helice of non-integral number of turns.

A helix can radiate in many modes, the three chief ones bein the normal, axial and conical modes. The particular mode dependent on the dimensions of the helix relative to the wav length of operation. Interest is generally centred in the axi mode of radiation, which is obtained when the length of on turn is of the order of one wavelength. As this length is varie one way or the other the width of the well-defined beam change until finally the radiation pattern splits and the helix goes int either the normal or conical mode of radiation (Fig. 1).

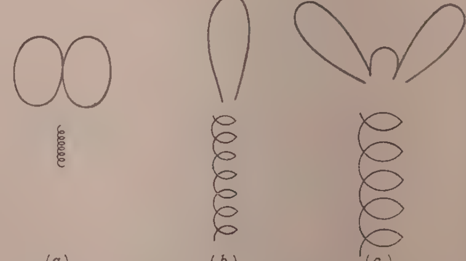


Fig. 1.—Three types of radiation pattern of a helix.
(a) Normal mode.
(b) Axial mode.
(c) Conical mode.

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
Mr. Jones is at the Royal Aircraft Establishment.

The normalized total radiation pattern for an axial-mode elix, assuming that the single-turn pattern is given by cos ϕ, expressed by

$$E = \sin\frac{\pi}{2n}\frac{\sin n\frac{\psi}{2}}{\sin\frac{\psi}{2}}\cos\phi \quad . \quad . \quad . \quad . \quad . \quad (1)$$

here n is the number of turns, and for the increased directivity ondition,

$$\psi = 2\pi\left[S_\lambda(1 - \cos\phi) + \frac{1}{2n}\right]$$

here S_λ is the spacing between turns in wavelengths.

The first factor in the expression is the normalizing factor, e. making the maximum value unity, and the second is the rray factor.

The beam width obtained with a single helix radiating in the xial mode on a flat ground plane is largely dependent on its xial length, i.e. the number of turns times the pitch, and can ary, at mid-band wavelength, from about 25° to 70° at half ower.

Kraus has developed the following empirical formulae:

Half-power beam width

$$\theta = \frac{52}{C_\lambda\sqrt{(nS_\lambda)}} \text{ degrees} \quad . \quad . \quad . \quad . \quad (2)$$

here D_λ = Diameter of helix, centre to centre, free-space wavelengths.

$C_\lambda = \pi D_\lambda$ = Circumference, wavelengths.

S_λ = Pitch, or spacing between turns, wavelengths.

n = Number of turns.

The pitch angle α is given by arc tan S_λ/C_λ.

The directivity of the helix compared with an isotropic source given by

$$D = 15C_\lambda^2 nS_\lambda \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The axial ratio, or ratio of the major to the minor axes of the olarization ellipse of the electric vector, is given by

$$AR = \frac{2n+1}{2n} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

The foregoing relations apply specifically to helices for $^{\circ} < \alpha < 15^{\circ}$, $\frac{3}{4} < C_\lambda < \frac{4}{3}$, and $n > 3$.

A minimum of two or three turns is required for good cirlarity. As the number of turns is reduced the axial ratio and e standing-wave ratio become quite high, and in an attempt determine the cause of this at the lower operating frequencies, oringer[4] made a systematic study of impedance and axial-ratio riation throughout the frequency region for axial-mode eration. He concluded that these high values were due to flection of the travelling wave at the end of the helix and that ey could be reduced by terminating the helix with a smaller lical coil. Axial radiation, without end-loading, is maintained en for helices as short as 1¼ turns, but the bandwidth over hich the axial ratio remains below a specified maximum also creases with a reduction in the turns. Fucci[5] has found that th short helices, either unloaded or loaded, the angle of the ajor axis of the polarization ellipse varies with frequency, but using a helix wound with strip material, the width of which tapered, the direction can be held substantially constant. single-turn helix is stated by Haycock and Ajika[6] to give a early-polarized axial beam, but circular polarization can be tained by introducing resistive end loading.

In general, the terminal impedance of helical aerials radiating in the axial mode is nearly a pure resistance, with a value between 100 and 200 ohms. Based on a large number of impedance measurements by Kraus, the terminal impedance of an axially-fed helix within the above specified limits, and mounted on a flat ground plane, is given by the empirical relation

$$R = 140C_\lambda \text{ ohms} \pm 20\% \quad . \quad . \quad . \quad . \quad (5)$$

Although a bandwidth of 2 : 1 is often claimed for a helical aerial, from the work detailed in this report, 1·7 : 1 seems to be a more reasonable claim as regards the polar diagram.

(3) S-BAND TESTS

(3.1) Single Helix Aerials, S-Band

In order to check whether Kraus's empirical formulae hold for the higher frequencies, a scaled-down helix consisting of six turns, of 3·1 cm diameter and 2·0 cm pitch, was wound with 0·113 in diameter copper wire and mounted on a 30 cm diameter flat metal plate. The radiation patterns were measured over a waveband of 8·8–13·2 cm, using the aerial as a receiver connected to a crystal detector, amplifier and recorder. Fig. 2 shows the

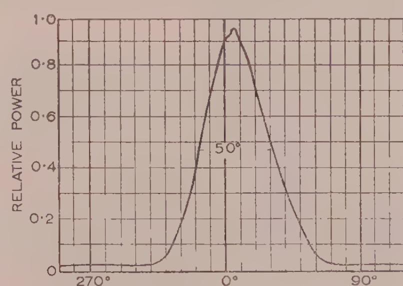


Fig. 2.—Radiation pattern of a 6-turn helix.
λ = 10·0 cm.

diagram recorded at 10 cm, giving a half-power beamwidth of 50° as compared with a calculated value of 47°. Good diagrams were obtained over the whole frequency range tested, and it is therefore assumed that the useful bandwidth could be extended some distance at each end of the 1·5 : 1 band covered by the tests. A circularity test made by rotating the helix aerial on its axis confirmed that the polarization was effectively circular.

A narrower beamwidth was explored with a 17-turn helix mounted in a 90° cone reflector of 10 cm aperture sunk flush on a metal plate. A half-power beamwidth of 27° was obtained at 10 cm, compared with a beamwidth of 25° as calculated from eqn. (2). This helix was approximately 45 cm long, and was too flexible for practical purposes.

(3.2) Four-Helix Array, S-Band

At this point it was of interest to attempt to reduce the axial length of the aerial while maintaining the narrow beamwidth, by constructing an array of helices. Four 6-turn helices were used, two stacked horizontally and two vertically, each being 15 cm in length, 3·2 cm in diameter and 2·5 cm in pitch. Each helix was fixed in a 90° cone of 10 cm aperture at the corner of a square of 10 cm side, i.e. λ spacing at mid-band, the cone apertures being set in flush with the surface of a 30 cm-diameter plate. The helices were fed in phase by coaxial leads, a tunable crystal detector was mounted directly on the back of the array, and the video-frequency output was fed through an amplifier to a recorder.

Good polar diagrams were obtained with this arrangement

from 8·9 to 15·9 cm, the latter not having reached the useful upper limit. The best-shaped diagram, with side lobes of less than 20 dB power, was obtained at 12·5 cm with a beam width of 30°, the axial power ratio being approximately 1·6 at this frequency. Azimuth diagrams were measured by rotating the helices in azimuth in either a horizontally or a vertically polarized field, and elevation diagrams similarly after turning both the helical and radiating aerials through 90°.

(3.3) Encapsulation of S-Band Helices

While this 4-helix array was more rigid than the single helix, the individual helices still required some support to prevent vibration and give protection. For this purpose each helix was surrounded by a resin-bonded glass-fabric tube 4·5 cm in diameter, which was filled with a foamed Sebalkyd resin. The resin adhered strongly to the cone and the glass-fabric tube, and gave adequate support. The foamed Sebalkyd resin had a density of 8 lb/ft^3, a permittivity of 1·16 and a loss factor of 0·17 dB/in at 3 cm. The presence of the combination of dielectrics around the helices did not appear to affect the performance of the aerial array appreciably over the S-band. Good polar diagrams were again obtained between 8·8 and 15·8 cm, the best being at 12·0 cm with a beam width of 27° and an axial power ratio of approximately 1·7.

(3.4) Investigation of Single Helix Parameters

During the course of the investigation a number of helices were wound with different diameters, or circumferences C_λ, and spacing between turns S_λ. The test results on these helices give data which extend those by Kraus, and are in some minor cases at variance with his conclusions.

For radiation in the axial-beam mode, and discounting the quality of axial ratio, Kraus implies that good polar patterns are not attainable with values of S_λ below 0·04 with a minimum value of $C_\lambda = 0·8$, where $\alpha \simeq 30°$, or above $S_\lambda = 0·43$ with a maximum value of $C_\lambda = 1·4$, where $\alpha \simeq 20°$. At both these extreme limits it is suggested that the bandwidth obtainable would be extremely small. Table 1 indicates that good patterns and bandwidth appear possible well outside these limits, a maximum value of $C_\lambda = 2·3$ being shown for helix 3, or mean values of $C_\lambda = 1·7$ and $S_\lambda = 0·065$ over the band tested.

The calculated beamwidths shown in the Table are based on Kraus's formula using a flat-plate reflector, although this was not expected to be applicable to these particular helix parameters.

Test results on helices 5 and 6 are included to show the feasibility of obtaining a broader beam pattern than is usually considered possible with a helix aerial.

The wide beams were obtained at abnormal values of C_λ and S_λ, and it is improbable that they would be maintained over a broad frequency band, or that the circularity would be particularly good. It should also be recorded that on a number of occasions when making repeat tests on 'unorthodox' helices certain discrepancies arose, such as the beam widths varying considerably, being, perhaps, only half of the maximum value previously obtained, although in some cases the patterns for both values were perfect beams without side lobes. These discrepancies were not investigated fully at the time, but nevertheless the results are included for completeness and as a possible useful future line to study.

A broad coverage might be more readily obtained under conditions which give a multi-lobed pattern if a perfectly shaped beam were not of first importance.

Springer[4] describes tests with an array of three helices which gave an azimuthal cover of well over 100°.

In the Table, under helix 7, it is shown how the calculated beamwidth varies over bandwidths of 1·7 : 1 and 2 : 1, namely 32° and 42° respectively, for optimum design at mid-band wavelength.

Table 1

SINGLE-HELIX TESTS ON FLAT-PLATE REFLECTOR

(Tests at 10·3 cm in Cone Reflector)

Helix number	Dimensions			Pitch angle α	Wavelength λ	C_λ	S_λ	Beam width		Remarks
	Diameter	Pitch	Turns					Calculated*	Test	
	cm	cm		deg	cm			deg	deg	
1	6·4	4·9	6	14	17·8	1·13	0·28	36	33	Very good pattern
					10·3	1·95	0·48	16	25	Very good pattern
2	6·4	1·8	2	5	17·8	1·13	0·10	102	66	Very good pattern
					10·3	1·95	0·18	43	36	8 dB power side lobe
3	6·4	0·75	6	2	17·8	1·13	0·04	92	70	Very good pattern
					10·3	1·95	0·07	40	40	7 dB power side lobe
					8·6	2·32	0·09	31	58	Good pattern
4	3·2	5·4	8	28	10·3	0·97	0·52	25	37	Very good pattern
5	3·3	3·6	2	19	17·8	0·58	0·20	142	98	Very good pattern
					10·3	1·01	0·35	61	51	Very good pattern
6	3·0	0·9	3	5	16·5	0·57	0·06	206	110	6 dB power dip at centre
					10·0	0·94	0·09	104	58	Very good pattern
7	3·2	0·25	6	14	13·3	0·75	0·19	65		Variation over 2 : 1 band
					12·6	0·80	0·20	59	Variation over 1·7 : 1 band	
					10·0	1·00	0·25	42		
					7·4	1·35	0·34	27		
					6·7	1·50	0·37	23		

* Calculated from formula not strictly applicable.

Helices 1, 2 and 3 are all of the same diameter but of differing umber of turns and spacing, and the test results show the ossibility of maintaining the beam width of a single helix onstant over the bandwidth by mechanically altering these arameters.

(4) X-BAND TESTS

(4.1) Single-Helix Aerials, X-Band

From the work so far carried out there appeared to be no eason for supposing that the helical aerial could not be scaled own further in size to operate in the 3 cm band, and the following sts showed that this is, in fact, possible.

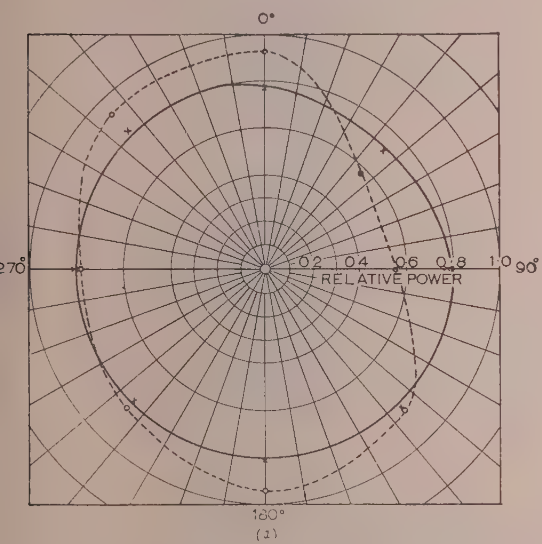
Several different sizes of helix wound with 0·06 in-diameter namelled wire were tested on a 20 cm-diameter earth plate, nd although good patterns were obtained, the beam widths were onsiderably wider than expected. Beam widths of 60° and 46° ere obtained for a 6-turn and a 12-turn helix, respectively.

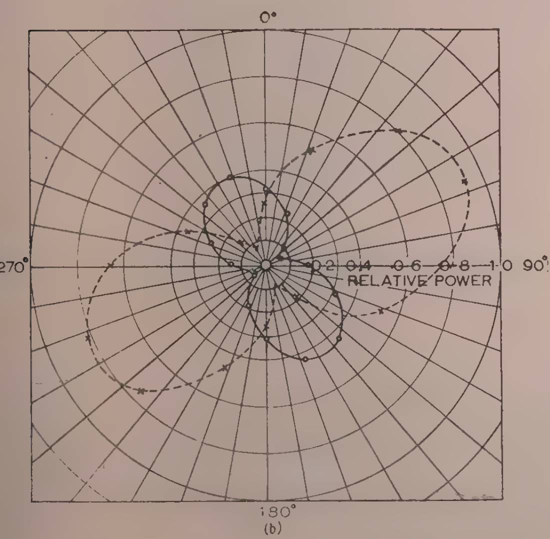
The use of a 90° cone reflector of λ aperture, set flush in a 20 cm-diameter plate, helped to reduce the side lobes on the X-band helix and also narrowed the beam somewhat compared with a 20 cm-diameter flat-plate reflector. On the S-band aerials the effect of the cone reflectors was hardly noticeable.

(4.2) Encapsulation of X-Band Helices

In order to give mechanical protection to the X-band helix it was encapsulated in foamed Sebalkyd resin similar to that used for the S-band aerial. On test, however, the results were very different, the polar diagram being badly distorted. Investigation showed that the foamed dielectric itself made little contribution to this adverse effect but that the distortion could be attributed to the resin bonding of the glass-fabric tube. Enclosing the helix with a lapping of unimpregnated glass fabric, a glass tube or a polystyrene tube brought about a considerable sharpening of the beam, from 46° to 25°, but did not split the diagram. A block of foamed resin with a hole in it was also slipped over the helix and sharpened the beam to a lesser extent.

It is seen that by a careful selection of the materials used it





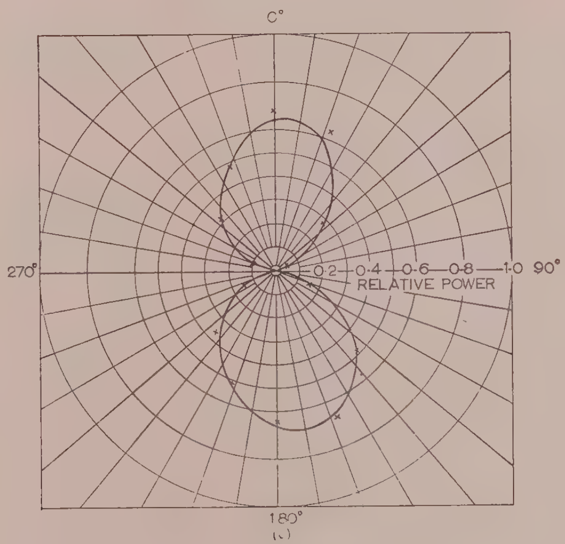


Fig. 3.—Circularity diagrams of rectangular helices.

(*a*) $\lambda = 10\cdot1$ cm.
——— Helix: 4·0 × 1·0 cm.
– – – – Helix: 4·5 × 0·75 cm.

(*b*) $\lambda = 9\cdot83$ cm.
Helix: 5·0 × 0·4 cm.
——— Vertical polarization.
– – – – Horizontal polarization.

(*c*) $\lambda = 9\cdot83$ cm.
Helix: 5·1 × 0·2 cm.

should be possible to obtain the desired polar diagram with the additional mechanical protection.

The bandwidth obtainable with the X-band helix was not investigated, most of the tests having been made at a fixed wavelength of 3·2 cm.

(4.3) Four-Helix Array, X-Band

An array of four helices similar to the S-band array, but scaled to cover the X-band, was made, each helix being approximately 0·9 cm in diameter and 0·75 cm in pitch and wound with 0·06 in-diameter enamelled wire. Good azimuth patterns giving beams of 20°–25° at 3·6 cm were obtained when receiving vertical polarization on an unencapsulated array, but some difficulty was found in radiating at wavelengths below 3·6 cm, probably owing to the poorly arranged feed system consisting of short unscreened leads from the aerials in a screened box. The output was taken through a coaxial plug and lead to a coaxial waveguide transition feeding a wide-band crystal detector.

(5) POSSIBILITIES OF WIDE-BAND LINEAR POLARIZATION

It was felt that the wide-band properties of helical aerials might, in some way, be used to provide wide-band linear radiation characteristics, and two possible methods were examined.

(5.1) Elliptical or Rectangular Helices

At first sight it would seem that the effect of making a circular helix elliptical by uniformly deforming it by a squashing action would be to vary the polarization from circular to elliptical and at the same time retain a good bandwidth. In the limit the zig-zag produced would be linearly polarized but would probably have lost in wide-band properties to a large extent.

Tests were initially made in the 300 Mc/s band on a 6-turn helix after three stages of deformation. Polar-diagram and circularity measurements were made, and at the maximum deformation, where the ratio of major to minor axis was approximately 6 : 1, the polar diagrams for both the horizontal and vertical components remained substantially unchanged and the axial ratio was not affected to any appreciable extent.

The considerable deformation of this particular helix was not extended at this stage, but further tests were made later on rectangular helices in the 10 cm band. Rectangular helices of different side-to-side ratio are very much easier to construct than elliptical helices, and should have similar properties.

Rectangular helices of side-to-side ratios of 1·0, 4·0, 6·0, 12·5, 14 and 25·5 were wound with six turns and tested in a coned counterpoise of 10 cm aperture mounted flush on a 30 cm-diameter plate. Fig. 3 shows circularity diagrams for helices of side-to-side ratios of 4·0 and 6·0, 12·5 and 25·5, while Table 2 gives details of the circularity or axial ratio.

Table 2

Helix dimensions	Side-to-side ratio	Axial power ratio
cm		
4·0 × 1·0	4·0	1·1
4·5 × 0·75	6·0	1·8
5·0 × 0·4	12·5	11·5
5·1 × 0·2	25·5	>35

The azimuth patterns of the vertical component of the 5·1 × 0·2 cm helix, side-to-side ratio of 25·5, are shown in Fig. 4 for wavelengths of 8·56, 9·83, 11·06 and 12·5 cm. It is seen that the beam width increases from 40° up to approximately

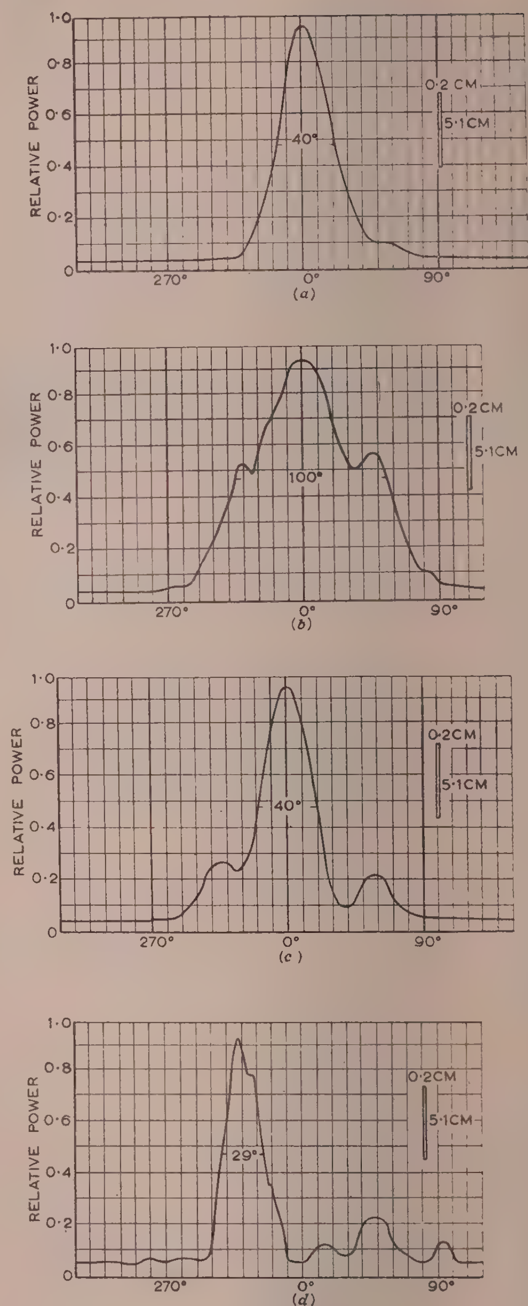


Fig. 4.—Radiation patterns of rectangular helix 5·1 × 0·2 cm.

(a) λ = 8·56 cm.
(b) λ = 9·83 cm.
(c) λ = 11·06 cm.
(d) λ = 12·50 cm.

100° and then decreases to approximately 30°, owing to a ris and fall in the side lobes.

It was hardly practicable to increase the side-to-side rati further without going to the limit of a flat zig-zag, and so 5 cm side, 2·5 cm pitch, 30° zig-zag was wound. This gav rather similar results to the 5·1 × 0·2 cm helix, but was teste

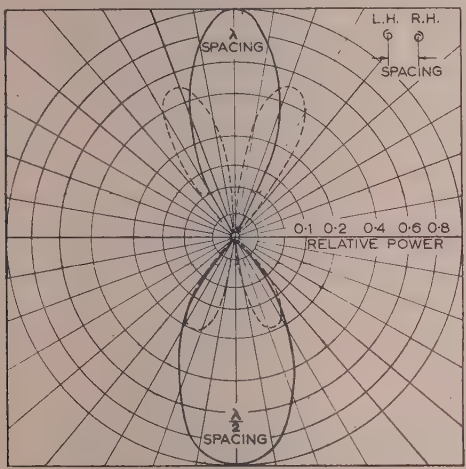
etween 7·7 and 13·1 cm. A square 90° zig-zag was tested at ne frequency, and it gave considerably greater gain than the 0° zig-zag.

From the limited tests made it appears that a useful bandwidth night be obtainable from a rectangular helix, but further work s necessary to establish the complete performance.

(5.2) Cancellation with Opposite-Sense Helices

If a right-hand and a left-hand helix are placed side by side nd fed in phase with equal power, the horizontal components f the fields are opposite in phase and cancel along the vertical lane through the axis of symmetry. At other points outside he plane the horizontal components cancel in varying degree, epending on the spacing of the two aerials.

Fig. 5 shows the calculated azimuth or horizontal-plane



ig. 5.—Calculated radiation patterns of left-hand and right-hand helices.

— — — Horizontal component. ———— Vertical component.

atterns for the two components with two different spacings of he aerials. Both the elevation and azimuth polar diagrams of he vertical component will be of similar form to the diagram f a single helix, but the beamwidths in the two planes will differ wing to the stacking factor in one plane only. Table 3 shows

Table 3

Spacing	Half-power beam width	
	Calculated	Test
	deg	deg
$1 \cdot 0\lambda$	26	26
$0 \cdot 7\lambda$	36	30
$0 \cdot 5\lambda$	40	35
$0 \cdot 3\lambda$	45	40
Single helix	48	50

he calculated and test values of beam width for the vertical omponent for different spacings of left- and right-hand 300 Mc/s elices, each wound with six turns to $C_\lambda = 1$ and $S_\lambda = 0 \cdot 22$, nd mounted on $\lambda/2$-diameter earth planes.

The test data confirm the calculations in showing that decreased pacing broadens the beam of the vertical component and reduces

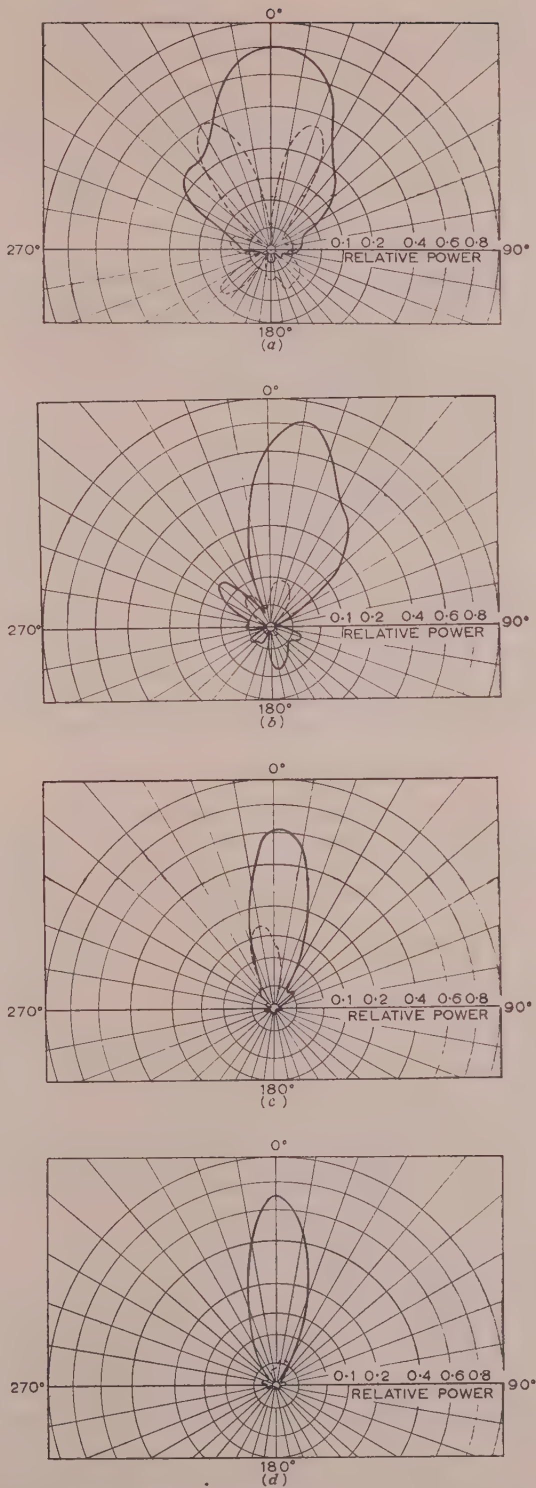


Fig. 6.—Radiated patterns of 4-helix array, left-hand and right-hand.

(a) Azimuth patterns at 550 Mc/s.
(b) Elevation patterns at 550 Mc/s.
(c) Azimuth patterns at 650 Mc/s.
(d) Elevation patterns at 650 Mc/s.
— — — — Horizontal component.
———— Vertical component.

the proportion of the horizontal component. A bandwidth of approximately 1·7 : 1 is maintained at $\lambda/2$ spacing.

If another pair of helices is stacked to form a 4-helix array a further plane of neutralization of horizontal polarization should result at right-angles to the first. The results of tests on a $\lambda/2$ spaced 4-helix array designed for 600 Mc/s are shown in Fig. 6. The azimuth polar diagrams at 550 Mc/s [see Fig. 6(*a*)] are seen to correspond more closely with those calculated for a 2-helix array (see Fig. 5), but the horizontal-component lobes at 20°–30° off the axis are larger than were expected, although they are a minimum along the axis, as was confirmed by a circularity test which showed linear polarization in this direction. The elevation polar diagrams [Fig. 6(*b*)] show a considerably reduced horizontal component, as do both the azimuth and elevation polar diagrams taken at 650 Mc/s [Figs. 6(*c*) and 6(*d*)].

Stacking of helices in this way does not seem to affect the bandwidth; a value of the order of 1·7 : 1 still appears to be possible at 300 and 600 Mc/s.

The above method of using left- and right-hand helices is briefly mentioned by Kraus, who also suggests the alternative possibility of connecting the two helices in series.

It can be shown theoretically that complete cancellation of one component in all planes results only when the spacing between a left-hand and a right-hand helix is reduced to zero, i.e. when they are coincident.

A rough attempt to check this was made by superimposing one of the left-hand and one of the right-hand 600 Mc/s helices already tested, but a circularity test showed an axial power ratio of 4 instead of linear polarization. It would be of interest to repeat the test with a more precisely made double-rotation helix.

(6) BAND-BROADENING DEVICES

(6.1) Tapered or Conical Helices

Although a normal helical aerial is essentially broad-band, approximately 1·7 : 1, attempts have been made to widen the band further by progressively varying the diameter of individual turns. Chatterjee[7] claims a useful range, for axial-mode radiation, of 120–450 Mc/s with a helix of 60 cm base diameter tapering to 20 cm in ten turns over an axial length of 112 cm and on a 90 cm-diameter earth plane, the feed point being at the apex. Good radiation patterns are shown for this helix between 150 and 450 Mc/s with a beamwidth of about 60°. The bandwidth is stated to be increased by adding turns, and the beamwidth is decreased, over a given frequency range, by keeping the diameter of the helix constant over a few turns at appropriate points. If the helix is fed from the base the bandwidth is smaller but the directivity greater, while feeding from the apex gives the largest useful bandwidth. Springer[4] also gives some impedance and radiation-pattern data on a tapered or expanding helix in the 1 000–3 000 Mc/s band.

With a view to the possibility of covering both the S- and X-bands on one aerial in this way, a tapered helix was made with four turns of 10 cm circumference tapering to four turns of 3 cm circumference; the total number of turns was 17 and the overall length 26 cm, with the pitch varying but S_λ maintained approximately constant at 0·25. The aerial was mounted with its base adjacent to a 30 cm-diameter flat plate or alternatively in a 10 cm aperture cone in a similar plate, and was fed from the base. Excellent polar diagrams were obtained in the cone reflector between 8·9 and 13·1 cm—the lowest test frequency—the beamwidth varying between 32° and 52°, respectively. At 8·6 cm the beam was split. An attempt to check the aerial at 3·2 cm gave no immediate indication of success and the work was temporarily suspended, but it is clear that there is considerable scope for further investigation with tapered helices.

(6.2) Other Possibilities

A few rather crude tests were made on a 3 cm test bench in the laboratory with no special support or turntable for the aerials, as compared with the previous tests which were conducted on well-equipped outdoor ranges.

A single 8-turn 10 cm helix and a single 6-turn 3 cm helix were mounted with centres approximately 6 cm apart on a 30 cm-diameter plate, with the feed points connected together at the back of the plate; the helices were not necessarily similarly orientated. With radiation from a 3 cm dipole and a 30 cm dish reflector the pick-up at 3·2 cm on the two helices appeared to be little different from that with the 3 cm helix alone. There was negligible pick-up of 3 cm radiation with the 10 cm helix only connected. The combination of the two helices was not tested for pick-up on 10 cm radiation.

The above experiment was varied by inserting the 3 cm helix concentrically inside the 10 cm helix and feeding them together. The result on 3 cm radiation was not as promising as before, since several lobes appeared to be present in the pattern. On disconnecting the feed to the 10 cm helix, a good beam with negligible side lobes was obtained. This arrangement would also appear to be worth further investigation.

For optimum design the diameter of a helix should vary with the wavelength of operation. A mechanical means of varying the spacing and the number of turns should not be too difficult to devise, but to vary the diameter would require considerably more ingenuity. It might be possible to construct a helix of springy material and fix the open end to an axially rotatable rod; the rod could also alter the spacing of the turns by an axial movement and provide a method of short-circuiting a few turns to maintain the beamwidth reasonably constant. An alternative arrangement would be to use a rectangular helix with telescopic sides.

Some of the properties of inductively end-loaded helices have been reported,[4,5] and claims are made of good impedance and pattern characteristics over extremely wide bands for such aerials.

(7) MISCELLANEOUS OBSERVATIONS

During tests on the 3 cm bench it was thought worth while to examine briefly the effect of placing metal rods and tubes coaxially with a helix. A 6-turn helix used with a cone reflector was covered with a closely fitting paper tube and the system tuned for maximum pick-up. A brass ring about $\lambda/4$ wide was slipped over the tube, and as it was moved along the helix the deflection of the galvanometer varied and presumably gave an indication of the current distribution. The deflection changed critically when the ring was over the turns nearest the cone, and appeared to be a maximum of 1·5–2 times the initial deflection with one edge of the ring in line with the cone aperture and a minimum with the ring pushed right inside the cone. A $\lambda/2$ wide ring was tried roughly and appeared to give reversed conditions to that of the $\lambda/4$ wide ring.

The insertion of a metal rod along the axis of the helix did not reduce the end-on deflection appreciably, but probably increased it to some extent when it was fully inserted.

In both the above cases it was hardly possible, without a conveniently movable support for the receiving helix, to determine the effect on the polar diagram; nevertheless the observations might be worth following up from the point of view of aerial matching or shuttering for crystal protection.

It is well known that the impedance of a helix aerial varies with the position of the end turn in relation to the earth plane, and this was confirmed by the very large changes in deflection obtained on making almost critical adjustments to the helix when close to the cone.

The diameter of wire used for the X-band helices tested was of he order of $0 \cdot 05\lambda$, and whilst it is known[8] that, in general, the ize of wire has little effect on the properties of the helix, it may e that this is getting rather large and its effect might be worth hecking at the higher frequencies.

(8) CONCLUSIONS

The non-critical nature of a helix aerial is clearly shown in the various tests described in the paper, particularly those with ectangular-shaped helices. It is found that considerable deformaion to a rectangular shape can take place before the axial ratio egins to increase rapidly. Even with a rectangular helix of ide-to-side ratio of $25 \cdot 5$ the radiation is still in the axial mode, ut the polarization has changed from circular to elliptical with n axial power ratio greater than 35; good polar patterns are till obtained over a wide band.

The above tests, together with those made with arrays of leftand and right-hand helices, indicates that a large amount of atitude is available in the design of an aerial to favour a particular lane of polarization to any required degree and yet to have comaratively good directivity and fair bandwidth.

Encapsulation of helical aerials in foamed dielectric material, n order to improve their rigidity, is found to be a satisfactory nd practical proposition provided that the materials are selected vith some care, particularly in the X-band.

Tests of single helices of varying diameter and pitch show that cceptable characteristics may be obtained with dimensions utside the limits suggested by Kraus. A beam width of 100°, t half-power points, is seen to be possible under certain limiting onditions.

A few tests were made on some tapered helices which are eputed to have wider bandwidths than the normal uniformly vound helix, but the full possibilities were not investigated. Other possible methods of increasing the bandwidth are quoted.

It is concluded that a helix aerial has such a variety of interesting possibilities that it deserves to be considered in some form for many applications where at first sight it might not appear to be suitable.

(9) ACKNOWLEDGMENTS

The author is indebted to Mr. W. T. Blackband for his helpful advice during the course of this work. Acknowledgment is made to the Controller, H.M. Stationery Office, for permission to publish the paper.

(10) REFERENCES

(1) KRAUS, J. D.: 'Antennas' (McGraw-Hill, New York, 1950).
(2) HANSEN, W. W., and WOODYARD, J. R.: 'A New Principle in Directional Antenna Design', *Proceedings of the Institute of Radio Engineers*, 1938, **26**, p. 333.
(3) KORNHAUSER, E. T.: 'Radiation Field of Helical Antennas with Sinusoidal Current', *Journal of Applied Physics*, 1951, **22**, p. 887.
(4) SPRINGER, P. W.: 'End Loaded and Expanding Helices as Broadband Circularly Polarized Radiators', *Proceedings of the National Electronics Conference*, 1949, **5**, p. 161.
(5) FUCCI, S.: 'Radiation Characteristics of Short Helices Radiating in the Axial Mode' (unpublished U.S. Navy Report).
(6) HAYCOCK, O. C., and AJIKA, J. S.: 'Radiation Characteristics of Helical Aerials of a Few Turns', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 989.
(7) CHATTERJEE, J. S.: 'Radiation Field of a Conical Helix', *Journal of Applied Physics*, 1953, **24**, p. 550.
(8) TICE, T. E., and KRAUS, J. D.: 'The Influence of Conductor Size on the Properties of Helical Beam Antennas', *Proceedings of the Institute of Radio Engineers*, 1949, **37**, p. 1296.
(9) GLASSER, O. J., and KRAUS, J. D.: 'Measured Impedances of Helical Beam Antennas', *Journal of Applied Physics*, 1948, **19**, p. 193.

DISCUSSION ON 'A SHORT MODERN REVIEW OF FUNDAMENTAL ELECTROMAGNETIC THEORY'

The paper (No. 1595), by MR. P. HAMMOND, *was published separately in December, 1953, and was republished, together with the London discussion and a number of written contributions, in July, 1954, in Vol. 101, Part 1, of the* PROCEEDINGS *(p. 147). A contribution by* MR. C. HARGREAVES *appeared in January, 1956, in Vol. 103 B (p. 22), after which it was considered that the correspondence on this subject should be closed. However, further contributions have been received—from* PROFESSOR E. G. CULLWICK, MR. HARGREAVES *and* MR. W. F. LOVERING—*and these have been carefully considered by the author in preparing this, his final reply to the discussion.*

Mr. P. Hammond (*in reply*): Mr. Hargreaves continues the discussion of the expanding rectangular circuit.* He suggests an rrangement by which the circuit would surround an iron core, he contention being that this would increase the $\frac{\partial B}{\partial t}$ and $\frac{d\phi}{dt}$ effects but would leave the $u \times B$ effect unaltered. Eqn. (22) would then give a different value for the e.m.f. from that given y eqn. (14).

Mr. Lovering takes up Mr. Hargreaves's problem. His conclusion is that eqns. (22) and (14) are always consistent, but he deprecates the division of the e.m.f. into a flux-cutting and a lux-changing part.

Prof. Cullwick questions the derivation of eqn. (22) from eqn. (14) for cases in which the circuit is not rigid. He prefers to derive the e.m.f. from the line integral of the Lorentz force $F = \mathscr{E} + u \times B$. He feels that problems containing a rotating disc or cylinder should be treated as a separate class, because the circuit in such cases contains no definite linear boundary.

All these contributions deserve a much more detailed reply than can be given in the space at my disposal.

The essence of Mr. Hargreaves's objection to the treatment of Faraday's law in my paper seems to lie in the transition from the equation

$$\oint \mathscr{E} . dl = -\frac{dt}{d}\iint B . da \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

to

$$\oint \mathscr{E} . dl = -\iint\left[\frac{\partial B}{\partial t} - \mathrm{curl}\,(u \times B)\right] . da \quad . \quad . \quad (22)$$

* Mr. Hargreaves wishes to point out that the third term in his eqn. (*a*) (see 103 B, . 22) should read

$$\frac{\sqrt{(L^2 + s^2)}}{L}$$

Mr. Hargreaves is prepared to accept eqn. (14) but not eqn. (22). As pointed out in the paper, the step from eqn. (14) to eqn. (22) is a mathematical one and involves no new physical facts except the universally accepted statement that div $\boldsymbol{B} = 0$ [eqn. (21)], i.e. that we have no experience of free magnetic poles. I do not share Prof. Cullwick's doubts as to the validity of Abraham and Becker's derivation of eqn. (22). Their treatment seems to be quite general, because the velocity $\boldsymbol{u}$ is given for each element of the circuit, which can therefore be deformed in any way whatsoever. In any case Prof. Cullwick arrives at the same results, although he prefers a different derivation.

The presence of an iron core would modify the magnetic field and attention would have to be paid to the induced surface polarity of the core, but Faraday's law as stated in eqn. (22) would continue to be applicable. Here at least Mr. Lovering and Prof. Cullwick agree with me. It cannot be pointed out too strongly that eqn. (22) is a mathematical statement which demands that the line integration of the left-hand side should exactly surround the area integration of the right-hand side. Current can only flow in conductors of finite cross-section. If, therefore, the line integration is carried out round a circuit that lies within the conductor material, it is essential to include the flux density within that material in the calculation. It is true that the additional area included in the integration may be small, but the flux density changes very rapidly in this small area and the contribution to the e.m.f. may be dominant. Since different paths of integration can be chosen which will give the same value for the line integral, Mr. Lovering is quite correct in pointing out that the flux cutting and threading terms are not unique. Nevertheless, in very many cases it is exceedingly helpful to separate the two effects.

Prof. Cullwick has consistently urged that the student's attention should be directed to the forces acting on electric charges. It can be seen from my paper that I entirely agree. But I am not convinced that it is wise in teaching to start with the Lorentz force and derive Faraday's law of induction from it. If the time comes when particle accelerators are more common in the laboratory than induction motors, I may change my mind. The sequence of instruction is clearly not very important. I do not feel that the homopolar generator is in a class by itself, because here again one is free to choose any convenient circuit for the integration of eqn. (22).

It is good to know that teachers will continue to debate these matters. The many contributions to this discussion have shown that fundamental theory is a very live issue, and the interest of the teacher cannot fail to stimulate interest in the student.

DISCUSSION ON 'THE NEW HIGH-FREQUENCY TRANSMITTING STATION AT RUGBY'*

Mr. K. L. Rao (*India: communicated*): It is generally known that the rhombic antennae have a large number of minor lobes and a certain amount of power is dissipated in the terminating resistor. The amount of power dissipated varies between 15% and 50% depending upon the size and construction of the rhombic. I suppose the authors considered the use of re-entrant networks for feeding back the power before deciding on the use of a dissipative terminating line. Re-entrant networks have been in use in some point-to-point communication networks using rhombics for the transmitting antenna. This arrangement, however, suffers from the disadvantage of being frequency sensitive, and thereby much of the aperiodic property of the rhombic is lost. The authors may be aware of the modified form of 'rat-race' which shows wide-band characteristics. Even though these have been successfully developed only in the v.h.f. and u.h.f. ranges, they could perhaps be extended to the h.f. range also.

Troubles due to minor lobes in the rhombic may be considerable, but there appears to be no mention in the paper of the measures taken to reduce them. The arrangement described in Section 6 makes no reference to this point. The technique of reducing the minor lobes appears to have been developed particularly in Australia. These could perhaps be used with advantage.

A time may soon come when the technique for reducing the minor lobes and the use of re-entrant networks for wide-band transmission will have developed. It would appear that the rhombic antenna will hold its own in all long-distance circuits. From past literature and installations it would appear that in the United Kingdom Kooman aerials have been preferred for communication/broadcast systems. The use of rhombics at the Rugby station is a happy departure from the above, particularly as in modern times site and aerial installation costs form a very important factor in station design. It is true that there is no measure of control over the horizontal and vertical lobes individually, and as such the rhombic may not be particularly useful in difficult circuits like the U.K.–New Zealand. However, when the coverage diagram is drawn, there is considerable spread as the number of hops increases. The deviations in long circuit paths may thus be covered up in actual practice.

Capt. C. F. Booth and **Mr. B. N. MacLarty** (*in reply*): Mr. Rao's comments are much appreciated, and in most cases he has supplied answers to his own questions. At Rugby the rhombic aerials were required to operate over a frequency range of $2\frac{1}{2}$: 1 or more, and this ruled out any re-entrant networks which were frequency sensitive. Further, although techniques for wide-band feeder systems at h.f. stations have been established, the application of re-entrant networks to the aerials is not at present in our programme and we would insist on the maintenance requirements of any future design being kept simple. The design of the rhombic aerials is such that the impedance of an aerial and its matching line has been kept fairly uniform to avoid unnecessary minor lobes. It is also worth while to emphasize the improvement in performance of the rhombic aerials which has been achieved by increasing the linear dimensions to approximately twice the values which had previously been used. This has increased the capital cost, and although, as Mr. Rao says, costs should be kept to a minimum, the additional expenditure has been found worth while.

* Booth, C. F., and MacLarty, B. N.: Paper No. 1903 R, October, 1955 (see 103 B, p. 263).

681.142:534:23

Paper No. 2194 R
Nov. 1956

ON THE MEASUREMENT OF ATTENUATION IN ULTRASONIC DELAY LINES

By M. REDWOOD, B.Sc.(Eng.), Graduate, and J. LAMB, Ph.D., Associate.

(*The paper was first received 9th May and in revised form 25th July*, 1956.)

SUMMARY

A theoretical and experimental study has been made of the effects of coupling films on the propagation of compressional waves from a transducer to a solid medium. In practice it has been found that 'wringing' the transducer to the specimen with oil as a coupling medium produces a film of non-uniform thickness. Although the variations in thickness are of the order of a wavelength of light, these variations are important at ultrasonic frequencies in the region of 50 Mc/s and above.

Conditions are described under which such films can lead to the propagation of a predominantly first-order mode in the specimen, resulting in an exponential decay of the amplitudes of successive reflections, with a consequent improved accuracy of attenuation measurement.

The work provides a greater understanding of the problems encountered in applications of delay lines.

(1) INTRODUCTION

In computers and radar-range measuring systems it is necessary to delay r.f. pulses with a minimum of distortion for periods ranging from a few microseconds to about one millisecond. Also it is often essential that the delay be constant and be known to an accuracy of the order of $\frac{1}{2}\%$. This may be achieved most conveniently by using an ultrasonic delay line. In principle the r.f. pulse is converted by a transducer into a pulse of mechanical oscillations which then travels through the delay medium. This may be a liquid, in which only compressional waves can be propagated, or a solid where, additionally, shear waves are possible. The delayed pulse is reconverted into an r.f. pulse by a second transducer; in some applications a single transducer is used both as transmitter and receiver.

Both solid and liquid delay lines are used in practice, a comprehensive bibliography of the literature being given in Reference 1. The performance of liquid delay lines has been studied in detail,[2,3,4] but for many purposes solid delay lines are preferred since they are more compact; their use, however, presents additional problems.[5,6] A particular study has been made by the authors of the attenuation of high-frequency compressional waves in fused quartz, which is the solid delay medium most commonly used. This work has relevance also to the techniques for measuring the absorption of sound waves in other solids. Valuable information concerning the fundamental physical properties of solid materials has been obtained from studies of this nature.[7–10]

Little attention has been given previously to the problems arising from the necessary use of a coupling medium between the transducer and the solid. It transpired from the authors' work that the properties of the coupling film are of great importance. A correlation has been found between the theoretical analysis of the behaviour of this film and experimental observation; this leads to a greater understanding of the performance of the system and to a more accurate determination of the attenuation in the solid than has hitherto been possible.

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Mr. Redwood and Dr. Lamb are in the Electrical Engineering Department, Imperial College of Science and Technology, University of London.

(2) PROPAGATION OF HIGH-FREQUENCY COMPRESSIONAL WAVES IN RODS OF CIRCULAR CROSS-SECTION

The experimental system is shown in Fig. 1. The oscillator produces pulses of 5 microsec duration at a frequency of between 30 and 100 Mc/s. These r.f. pulses [500 per second, of approximately 200 volts (peak-to-peak)] are applied to the faces of an X-cut quartz-crystal transducer on which evaporated gold or aluminium electrodes have been formed. The fundamental resonant frequency of the transducer is 10, 12·5 or 15 Mc/s, and it operates at its 3rd, 5th or 7th harmonic. A thin film of oil is used as a mechanical coupling between the transducer and the specimen of fused quartz. Each pulse undergoes multiple reflections in the specimen, and the received signals are displayed on the oscilloscope after amplification and detection.

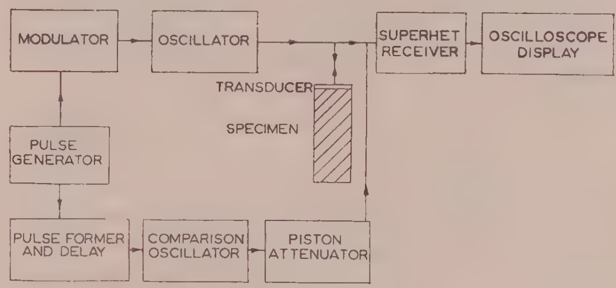


Fig. 1.—Apparatus for the study of propagation in solid delay lines.

A second r.f. pulse is formed by the 'comparison' oscillator and fed to the receiver through a piston attenuator; this pulse occurs at a controlled time interval after the main transmitted pulse. It is viewed on the oscilloscope alongside the reflections in the specimen, and enables the difference in amplitude between any two pulses to be determined.

Ideally the amplitude of the waves reflected in the specimen should decrease exponentially with the distance travelled, e.g. Figs. 9(*c*) and 9(*d*). The rate of decay gives the total loss, which includes the losses upon reflection at the ends in addition to the intrinsic absorption in the medium. The loss occurring on reflection at the free end of the specimen is usually extremely small, but that at the transducer end may, under certain conditions, be determined experimentally, as described in Section 4. It is assumed that propagation takes place at a single frequency only: this is not true in practice when pulses are employed but is a good approximation for pulses containing a large number of cycles.

In the present frequency range the wavelength λ_S in the specimen is much less than either the diameter of the transducer, d_T, or that of the specimen, d_S. For example, at a frequency of 60 Mc/s, $\lambda_s = 0{\cdot}01$ cm, while d_T and d_S are normally of the order of 1 cm. Several previous workers appear to have used transducers of smaller diameter than the specimen ($d_T < d_S$). In this case an additional loss occurs owing to the spreading of the 'beam' of energy leaving the transducer. For large values of

d_T/λ_S the transducer may be regarded as producing a directional 'beam' of energy, for which the semi-angle is approximately λ_s/d_T. This angle is generally small (about $\frac{1}{2}°$ at 60 Mc/s if $d_T = 1$ cm), but the resulting decrease in intensity cannot be neglected when the intrinsic loss in the specimen is small. It is not possible to determine experimentally the apparent loss due to beam spreading, nor is the system amenable to theoretical calculation.

In order to avoid this unknown error due to beam spreading a 'guided-wave' technique was chosen, with $d_T = d_S$, the transducer completely covering the end of the rod. The theory of the propagation under these conditions has been studed by McSkimin,[5] and is closely analogous to the propagation in a mercury-filled cylinder.[3] At any one frequency it is possible to propagate a number of modes which have slightly different phase velocities, the effect being very similar to the propagation of E (or TM) modes in a circular waveguide. Interference results when conditions are such that two or more modes are excited simultaneously (as usually occurs). The degree of cancellation is a function of the distance travelled; the observed decay then has the form illustrated in Fig. 2.

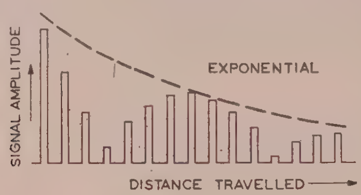


Fig. 2.—Decay pattern observed when there is interference between modes.

Mason and McSkimin[11] found it possible to produce a wave consisting almost entirely of one mode (the first-order mode) by using a shaped electrode to produce a suitable distribution of the electric field strength across the surface of the transducer. In the authors' experimental work, however, it was found that uneven decays similar to Fig. 2 occurred both with and without a shaped electrode. Moreover, as the frequency was altered, the change in position of the pulse of minimum height was much greater than could be accounted for by theoretical consideration of interference between modes in the specimen.

Further investigation led to the conclusion that the predominant cause of the uneven decay is the inherent variation in the thickness of the coupling film across the face of the transducer.

The customary method of making the coupling film is to 'wring' the polished transducer to the polished and plated face of the specimen with a little oil, thus producing a film thickness of the order of one wavelength of light. It is possible to investigate the form of such a film optically, since the transducer is transparent to light, and interference patterns caused by the film may be observed. Examination of these patterns showed that, in general, the thickness was about 3 000Å near the perimeter of the transducer, but three or four times greater near the centre. Coupling films of this nature were reproducible and were as thin as could be obtained without damaging the plated electrodes.

Theoretical consideration of the effects of such non-uniform films has provided explanations for many of the peculiarities observed in the experimental work, particularly the non-exponential decay patterns.

(3) THE EFFECTS OF THIN COUPLING FILMS

The following Sections deal with the effect of the coupling film on the amplitude and phase of the pressure wave in the specimen when this is reflected from, or transmitted through, this film. In the process of expounding the theory it will be convenient at each stage to make comparisons with experimental observation.

The mechanical and acoustical systems will be represented by equivalent electrical transmission lines, assuming plane-wave propagation and taking force as being analogous to voltage and velocity to current.[12]

(3.1) Equivalent Circuits of an X-Cut Quartz Transducer

The equivalent circuit of a vibrating X-cut quartz crystal with arbitrary loading on each face has been developed by Mason[13] and is shown in Fig. 3. The network elements are lumped

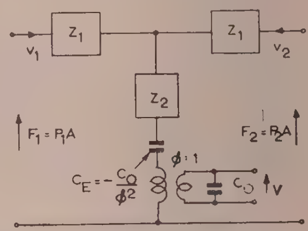


Fig. 3.—Equivalent circuit of a thickness-vibrating quartz transducer.

F = Force in the medium adjoining the transducer.
p = Pressure in the medium adjoining the transducer.
v = Velocity in the medium adjoining the transducer.

$$Z_1 = jZ_T \tan \frac{\omega l_T}{2c_T}$$

$$Z_2 = -\frac{jZ_T}{\sin \frac{\omega l_T}{c_T}}$$

l_T = Transducer thickness.
c_T = Velocity of compressional waves in the transducer = $\sqrt{(c_{11}/\rho_T)}$.
$\omega = 2\pi \times$ frequency of operation.
Z_T = Mechanical impedance of quartz = $A\sqrt{(\rho_T c_{11})}$.
ρ_T = Density of quartz.
A = Surface area of transducer.
c_{11} = Elastic constant of quartz.
$\phi = \frac{D\kappa A}{4\pi l_T}$ and converts mechanical units to electrical and vice versa.
V = Applied voltage.
I = Current.
κ = Dielectric constant of quartz.
D = Piezo-electric constant of quartz.
$C_0 = \frac{\kappa A}{4\pi l_T}$ = Electrostatic capacitance.

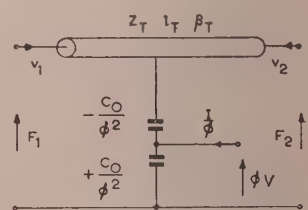


Fig. 4.—Equivalent circuit of transducer.

Z_T = Characteristic mechanical impedance of line representing quartz.
l_T = Length of line representing quartz.
$\beta_T = \omega/c_T$.
c_T = Velocity of compressional waves in quartz.

systems, but for present purposes it is convenient to replace part of the circuit of Fig. 3 by an exactly equivalent length of transmission line as shown in Fig. 4. It is also convenient to have an approximate equivalent circuit for a crystal with one face free, and this is shown in Fig. 5. The latter was also derived by Mason[13] (from the circuit of Fig. 3) and holds at frequencies near each odd harmonic of the fundamental transducer resonance.

(3.2) Reflections at the Transducer

The equivalent circuit of Fig. 4 may be used to investigate the changes in amplitude and phase which occur when a reflection takes place at the transducer end of the specimen—a condition represented by Fig. 6. The pulse arrives from the specimen

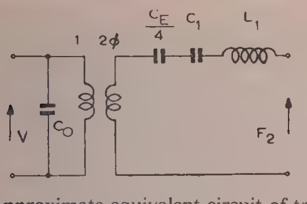


Fig. 5.—Approximate equivalent circuit of transducer.

$$C_1 = \frac{2l_T}{\pi^2 c_{11} A} \times \frac{1}{n^2}$$
$$L_1 = \frac{\rho A l_T}{2}$$
n = Order of harmonic.

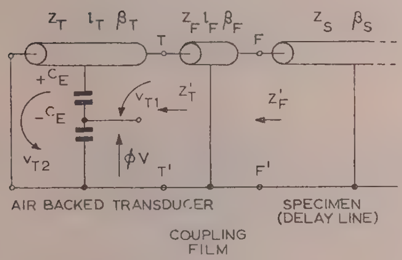


Fig. 6.—Equivalent circuit for line, transducer and film.
v_{T1} and v_{T2} = Particle velocities at the transducer faces. Z, β, and l are line parameters.

(delay line) of characteristic impedance Z_S, which terminates in a length l_F of transmission line of characteristic impedance Z_F representing the coupling film and a further length, l_T, Z_T, representing the transducer. One face of the transducer is assumed to be free, and electrical open-circuit conditions are assumed in order to simplify the calculations. The two capacitances in the equivalent circuit of the crystal may be ignored as they are equal and of opposite sign. All elements are assumed to have no loss.

(3.2.1) **Terminating Impedance.**

The mechanical input impedance Z'_T of the transducer, as seen from the points TT′ (Fig. 6) is given by transmission-line theory:

$$Z'_T = jZ_T \tan \theta_T \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Here $\theta_T = \beta_T l_T$, where l_T is the transducer thickness; $\beta_T = \omega / C_T$, C_T is the velocity of compressional waves in the transducer; and $\omega = 2\pi \times$ frequency.

The mechanical impedance Z'_F of film and transducer together, seen from FF′ is given by

$$Z_F' = Z_F \frac{jZ_T \tan \theta_T + jZ_F \tan \theta_F}{Z_F - Z_T \tan \theta_T \tan \theta_F} \quad . \quad . \quad . \quad . \quad (2)$$

Here the symbols are as previously described, the suffix F referring to the film.

The variation of Z'_F with frequency is plotted in Fig. 7. The particular curves shown illustrate the behaviour between 60 and 80 Mc/s of a 1 cm² quartz transducer of fundamental frequency 10 Mc/s, coupled through a film of oil to a fused-quartz line. The characteristic impedance of the transducer, Z_T, is assumed to be equal to that of the specimen, Z_S, and the impedance, Z_F, of the oil film is taken to be $\frac{1}{10}Z_S$. Four thicknesses of film are considered, referred to as 0°, 5°, 10° and 20° films, representing line lengths (expressed as fractions of a wavelength) of 0, $\frac{5}{360}\lambda_{70}$, $\frac{10}{360}\lambda_{70}$ and $\frac{20}{360}\lambda_{70}$, where λ_{70} is the wavelength in oil at 70 Mc/s. The actual thicknesses of these films are approximately 0, 3 000 Å, 6 000 Å, and 12 000 Å.

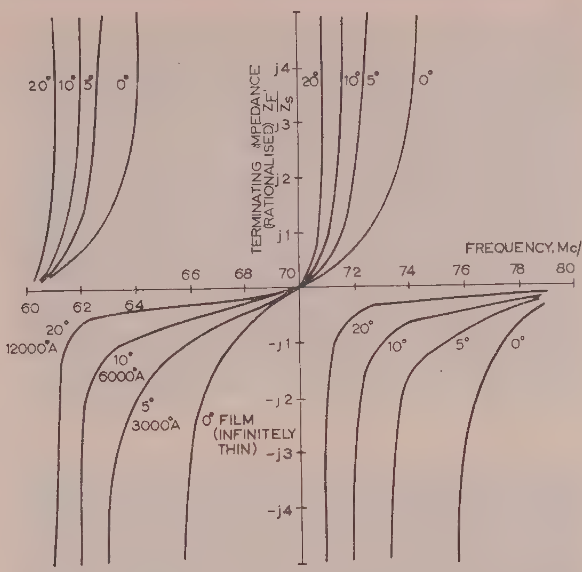


Fig. 7.—Calculated terminating impedance (film and transducer).

$$Z_F' = jZ_F \frac{Z_T \tan \theta_T + Z_F \tan \theta_F}{Z_F - Z_T \tan \theta_T \tan \theta_F}$$

Transducer of 10 Mc/s fundamental resonance; area = 1 cm²; oil film $Z_F = \frac{1}{10} Z_S$; $Z_S = Z_T$.

Fig. 7 shows that one effect of the film is to alter the frequency at which the terminating impedance Z'_F is infinite; the thicker the film, the nearer to 70 Mc/s does this infinity occur. (It may be noted that a quarter-wavelength film is necessary to cause the terminating impedance to appear infinite exactly at the harmonic of 70 Mc/s.)

(3.2.2) **Reflection Coefficient.**

Of great importance is the behaviour of the complex reflection coefficient R, which is defined as the ratio of the reflected wave to the incident wave:

$$R = \frac{Z'_F - Z_S}{Z'_F + Z_S} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

In Appendix 8.1 it is shown that $|R| = 1$ (since Z'_F is purely reactive), and that the phase ϕ_R of the reflection coefficient is given by

$$\phi_R = \pi - 2 \arctan a \quad . \quad . \quad . \quad . \quad (4)$$

where $a = Z'_F / jZ_S$, and is a function of frequency.

Fig. 8 shows the variation of ϕ_R with frequency for various thicknesses of coupling films. (Calculations for Figs. 7 and 8 were greatly simplified by use of a Smith chart.)

(3.2.3) **Reflections at a Non-uniform Film.**

It has been stated previously that the coupling-film thickness varies between about 3 000 Å (5°) and 12 000 Å (20°) across the face of the transducer. For this reason the phase shift that occurs when a pulse is reflected at the interface between coupling film and specimen is not constant over the wavefront, and loss of signal amplitude results.

For example, at 69 Mc/s the difference in phase between those parts of the wavefront reflected from the centre and perimeter of the transducer is about 15°, while at 70 Mc/s it is very small, and at 71 Mc/s it is about 100°. This phase difference occurs at each reflection, so that the phase differences across the wavefront for successive pulses arriving at the transducer will be 0°, 15°,

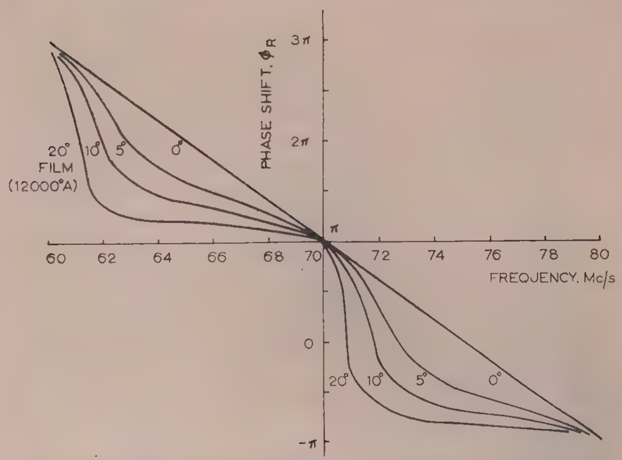


Fig. 8.—Calculated phase shift occurring upon reflection at transducer end of specimen.

$$\phi_R = \pi - 2 \arc \tan a$$

$$a = \frac{Z_F}{Z_S} \frac{Z_T \tan \theta_T + Z_F \tan \theta_F}{Z_F - Z_T \tan \theta_T \tan \theta_F}$$

$$Z_S = Z_T = 10 Z_F$$

30°, 45°, etc., at 69 Mc/s, and at 71 Mc/s 0°, 100°, 200°, 300°, etc. The transducer acts as an integrating device, and in order to determine the phase difference across the wavefront which is required for complete cancellation of the signal it is necessary to take into account the pressure variation over the wavefront and the variation in the thickness of the film. The exact form of these variations is in most cases unknown, but it is shown in Appendix 8.2 that, with certain assumptions approximating to practical conditions, a complete cancellation occurs with a phase difference of 360° between the centre and outside of the wavefront. Experimental observation indicates that this is a reasonable estimate.

The experimental findings are illustrated by the photographs of Figs. 9(*a*)–9(*f*), which show decay patterns found at 67·6, 68·8, 70·2, 71·3 and 73·3 Mc/s.

It was found that the position (in the decay pattern) of the first pulse showing maximum cancellation altered as the frequency was increased from 65 to 75 Mc/s. As resonance (70 Mc/s) was approached, maximum cancellation occurred after an increasing number of reflections had taken place [cf. Figs. 9(*a*) and 9(*b*)] until near 70 Mc/s no minimum was detectable [Fig. 9(*c*) and 9(*d*)].

These observations agree qualitatively with the theoretical considerations that led to Fig.8, which shows the phase difference across the wavefront to be decreasing as resonance is approached.

At 70·2 Mc/s the decay was almost perfectly exponential [Figs. 9(*e*) and 9(*d*)].

A further slight increase in frequency caused a very rapid change in the decay pattern; at 71·3 Mc/s the fourth and fifth pulses show maximum interference [Fig. 9(*e*)].

Theoretical considerations show that the phase differences that occur at frequencies between 1 and 2 Mc/s above resonance are very much greater than those which occur between 1 and 2 Mc/s below resonance. The experimental observations, therefore, agree qualitatively with the theory.

It is predicted from theory that, as the frequency is increased above about 71·5 Mc/s, the phase difference will diminish (see Fig. 8) and that the pulse showing a minimum will not appear until more reflections have occurred. This also agrees closely with experimental observation [cf. Figs. 9(*e*) and 9(*f*)].

(3.3) The Effect of a Film on Transmitted Pulses

(3.3.1) Phase Change due to a Coupling Film.

A phase change occurs also when a signal passes *through* the film; this happens twice, once when the signal is transmitted initially and once when it is received, and it may be shown that the phase changes on transmission and reception are identical. The phase change on transmission, determined by the method of Appendix 8.3, is equal to $\frac{1}{2}(\phi_R - \pi)$, where ϕ_R is the phase shift on reflection as plotted in Fig. 8. Once again a phase difference across the wavefront results, but this is the same for every pulse arriving at the transducer. A further phase difference is thus added to the cumulative one caused by multiple reflection, although the effect of the latter is usually predominant.

(3.3.2) The Effect of a Film on Signal Amplitude.

The effect of a coupling film on the amplitude of the pressure wave in the rod has also to be considered. This has been determined using the method of Appendix 8.4, and is shown graphically in Fig. 10. It is found that the effect of a uniform coupling film of finite thickness leads to an increased pressure amplitude in the rod at a frequency slightly above resonance (the transducer both transmitting and receiving with greater efficiency) and also to a decreased bandwidth, the system acting as a transforming band-pass filter. This result, in a slightly different form, has been described by previous workers.[11]

Since, in practice, the coupling film is thicker in the centre than at its periphery, at a frequency slightly above resonance the maximum pressure in the radiated wave will occur near the centre of the rod, where the film is thickest. It is considered that, under these conditions, the wave propagated in the delay medium consists predominantly of the first-order mode. Theoretical considerations show that the pressure distribution required to produce this first-order mode is a combination of two Bessel functions and may be considered to fall in a roughly linear manner from a maximum at the centre of the rod to zero at the outside.[5] This single mode of propagation is confirmed experimentally by the fact that, at a frequency slightly above resonance, good exponential decays are observed [Figs. 9(*c*) and 9(*d*)].

The two factors due to the presence of the coupling film, namely greater amplitude and decreased bandwidth, are believed also to account for an observed deterioration in pulse shape occurring just above the harmonic of the resonance frequency, and an observed maximum signal strength occurring simultaneously with the deterioration.

(3.4) Experimental Observation of Interference between Modes

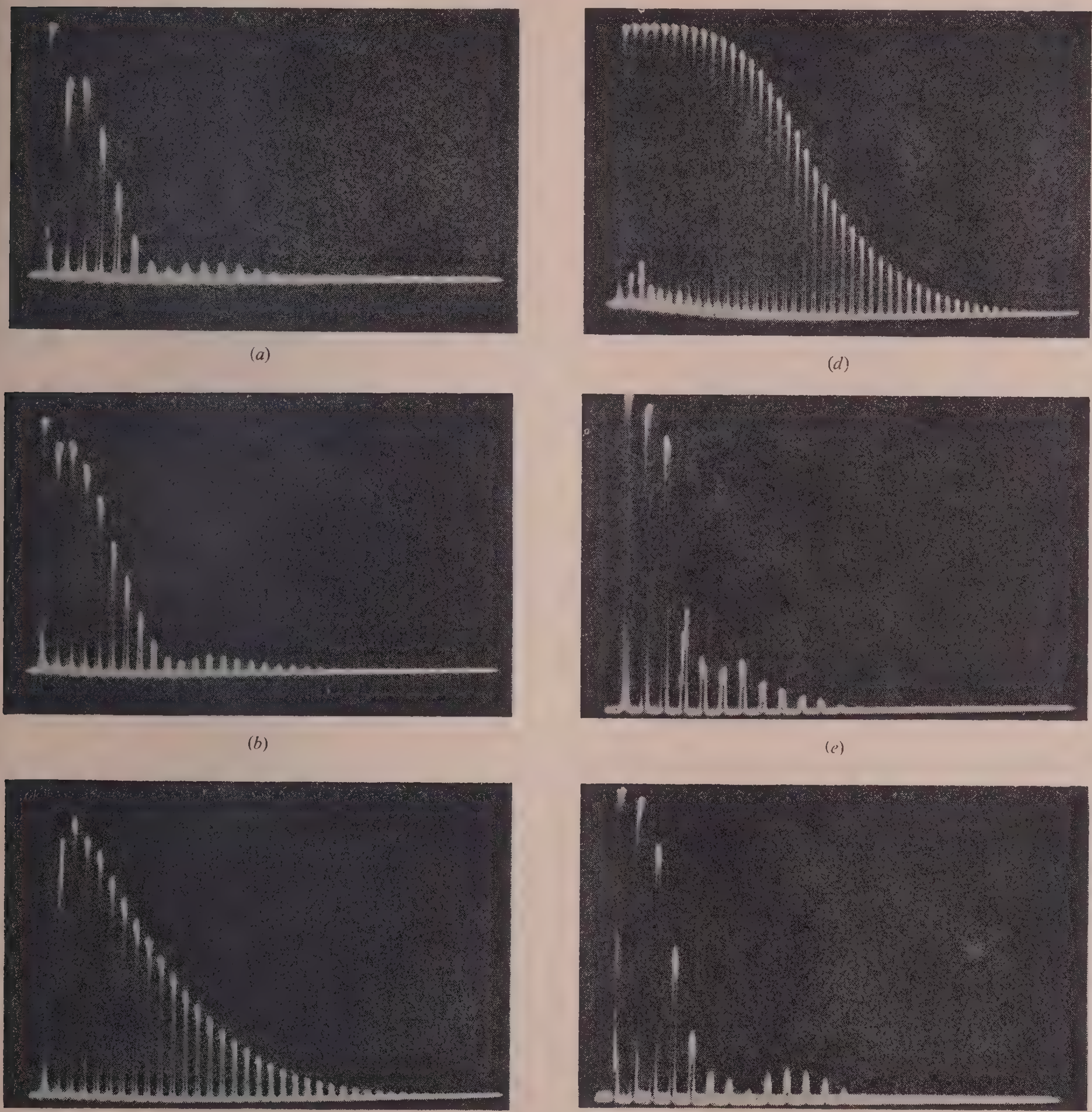
In order to investigate the effect of mode interference as a separate phenomenon, an unplated transducer was 'wrung' as tightly as possible to an unplated rod until no optical interference fringes were visible. Uneven decays occurred at all frequencies, the position of the first minimum pulse changing only slightly with frequency (the 7th pulse being a minimum at 45 Mc/s and the 8th at 55 Mc/s).

The velocities of propagation, v, of the different modes are given by

$$v^2 = \frac{v_0^2}{1 - \left(\frac{K}{\omega}\right)^2 v_0^2} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where v_0 is the free-space velocity of sound in the material and K is a constant depending on the order of the mode and the radius a of the specimen.

v may be evaluated with sufficient accuracy for present purposes



(a) (d)

(b) (e)

(c) (f)

Fig. 9.—Decay patterns in fused quartz rod.

Specimen of fused quartz 1·4 cm diameter × 5·9 cm long. Pulse length, 5 microsec. First pulse is the transmitter pulse, arriving by direct electrical path. In some photo-
raphs the first reflected pulse arrived before the receiver had fully recovered.

(a) $f = 67{\cdot}6$ Mc/s. 7th pulse a minimum.
(b) $f = 68{\cdot}8$ Mc/s. 10th pulse a minimum.
(c) $f = 70{\cdot}2$ Mc/s. Good exponential.

(d) $f = 70{\cdot}2$ Mc/s. Continuation of decay of Fig. 9(c) (increasing the receiver gain caused the first 8 pulses to saturate the receiver and pulses 8–16 to receive non-linear amplification). 45 reflections are visible.
(e) $f = 71{\cdot}3$ Mc/s. 4th and 5th pulses a minimum.
(f) $f = 73{\cdot}3$ Mc/s. 7th pulse a minimum.

y assuming Ka to be approximately equal to the roots of the Bessel function[5] $J_0(Ka)$. For the conditions under which the resent experiments were made, this leads to a difference in elocity between the first two modes of 35 cm/s at 45 Mc/s and 3 cm/s at 55 Mc/s. For 180° phase difference between these two modes (which, it is assumed, are predominant) 110 and 135 cm of path lengths are required, respectively, i.e. about nine reflections at each frequency.

The experimental observations quoted agree closely with these figures.

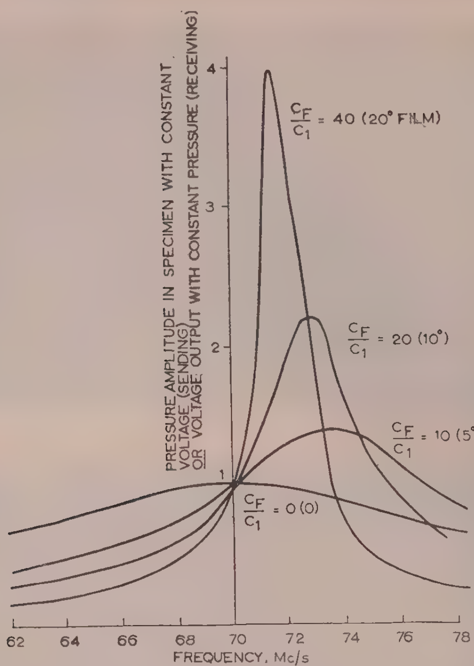


Fig. 10.—The calculated effect of a coupling film on the efficiency of the transducer.

Pressure amplitude in the rod

$$\propto \frac{\text{Voltage across transducer}}{\sqrt{\left\{\left[1 + \frac{C_F}{C_1}\left(1 - \frac{f^2}{f_0^2}\right)\right]^2 + \left[\frac{1}{\omega C_1 Z_S}\left(1 - \frac{f^2}{f_0^2}\right)\right]^2\right\}}}$$

(4) MEASUREMENT OF ATTENUATION

Exponential decays, obtained by the methods just described, are being used to investigate the properties of several low-loss materials, including fused quartz. This 'decay' measuring technique is particularly well adapted to the determination of absorption in low-loss specimens, owing to the very large effective path length [the 40 reflections of Figs. 9(*c*) and 9(*d*) represent nearly 500 cm of effective path length].

Fig. 11 shows a typical exponential decay pattern for fused quartz at 45 Mc/s plotted on a decibel scale. The specimen was of Thermal Syndicate OG fused quartz 1·4 cm diameter × 5·9 cm long, with the ends polished and parallel to within 30″. The first few points lie off the main curve; the reason for this is thought to be that the mode of propagation is not set up completely until three or four reflections have taken place. It will also be seen that a slight, regular deviation from linearity occurred; this may have been caused by mode interference, or by interference across the wavefront owing to multiple reflection.

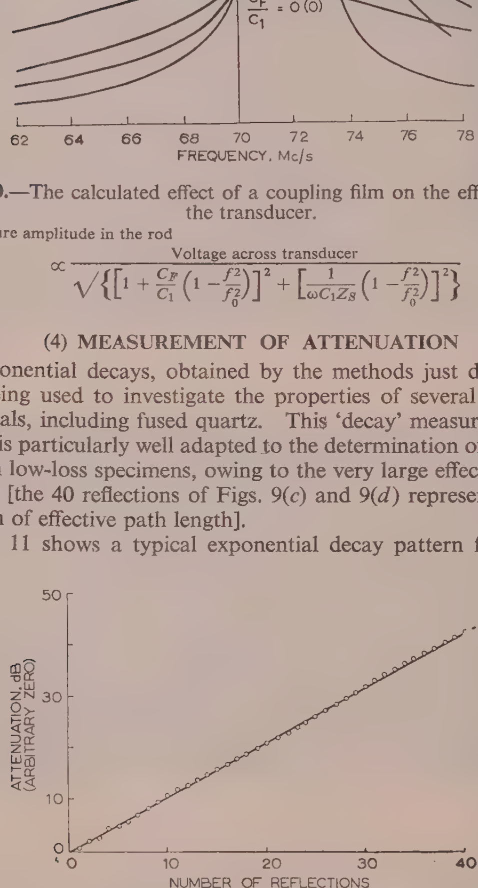


Fig. 11.—Plot of attenuation against distance travelled in specimen.
Specimen of fused quartz 5·9 cm long × 1·4 cm diameter. $f = 45$ Mc/s.
Attenuation 1·06 dB/double trip of 11·8 cm.

The slope gives the attenuation as 1·06 dB per double trip (11·8 cm) at 45 Mc/s. This figure includes one loss on reflection at the transducer, owing to energy taken by the electrical circuit and, possibly, losses in the oil film. The reflection loss was determined by using a second identical transducer, placed on the end of the specimen which had previously been free. Two more measurements of attenuation were then made, transmitting from the second transducer—the first measurement with the original crystal still in position and the second with the original crystal removed. The difference between these two sets of attenuation measurement gave the reflection loss at the original transducer (in this particular instance equal to 0·96 dB), which was then subtracted from the result given by the original set of measurements to obtain the intrinsic attenuation in the medium. (This method has been developed independently by McSkimin.[5])

Preliminary results show that the absorption in fused quartz at 45 Mc/s is less than 0·01 dB/cm, representing a Q-factor of greater than 200 000. It is necessary, however, to apply a correction factor to allow for the conversion of compressional waves to shear waves which occurs at the side walls of the specimen. An expression for this has been derived from theoretical considerations by McSkimin.[5] In the case just cited the application of the correction factor would raise the Q-factor to above 300 000. Further measurements are at present being made at higher frequencies, where the loss due to conversion is less and the intrinsic absorption greater.

(5) CONCLUSIONS

Propagation of mechanical compressional waves through solid media is used in delay lines and in the measurement of the fundamental properties of solid materials. For such applications it is necessary to employ some form of mechanical coupling between the transducer and the solid. In the work described in the paper, oil films have been used to provide this coupling.

Although the faces of both the specimen and the transducer are optically polished, the transducer is sufficiently thin and flexible to deform during the 'wringing' procedure (a transducer of 10 Mc/s fundamental frequency is approximately 0·01 in thick). Films made in this way are three to four times thicker at the centre than at the periphery of the transducer. It has been shown theoretically that the maximum pressure on the end face of the specimen occurs at the centre where the film is thickest. This results in the excitation of a predominating first-order mode in the solid (with the consequent elimination of interference between possible modes) and a good exponential decay of successive reflections. In order to achieve this effect it is necessary to operate the transducer at a frequency slightly above a harmonic of its fundamental resonance. Using this technique it is possible to measure the attenuation of the pulses propagated through the specimen with an accuracy of $\pm 1\%$; after allowing for the losses occurring on reflection at the transducer this leads to a determination of the intrinsic loss in the solid material.

(6) ACKNOWLEDGMENTS

The work described in the paper has received the financial support of the Worshipful Company of Clothworkers, to whom the authors' thanks are due. One of the authors (M. R.) wishes to acknowledge financial assistance through a University of London Postgraduate Studentship (1954–55), a Supplementary Award from the Department of Scientific and Industrial Research

(1954–55) and the Oliver Lodge Scholarship of The Institution of Electrical Engineers (1955–56).

The authors are also especially grateful to Dr. W. P. Mason and Mr. H. J. McSkimin of the Bell Telephone Research Laboratories for a personal communication concerning their work on mode propagation in solid rods. Thanks are due also to Dr. J. R. Barker of the Imperial College of Science and Technology, for helpful discussion on the equivalent circuit representation of quartz transducers, and to a referee for suggesting improvements in the expressions for phase shift in Appendices 8.1 and 8.3.

(7) REFERENCES

(1) FAGEN, M. D.: *Transactions of the Institute of Radio Engineers* (Professional Group on Ultrasonics Engineering), 1954, p. 3.
(2) HUNTINGTON, H. B.: *Journal of the Acoustical Society of America*, 1948, **20**, p. 424.
(3) MCSKIMIN, H. J.: *ibid.*, 1948, **20**, p. 418.
(4) DAIRIKI, S., LAWRENCE, T. E., and MAPLETON, R. A.: *ibid.*, 1953, **25**, p. 841.
(5) MCSKIMIN, H. J.: *ibid.*, 1956, **28**, p. 484.
(6) ARENBERG, D. L.: *ibid.*, 1948, **20**, p. 1.
(7) BOMMEL, H. E.: *Physical Review*, 1954, **96**, p. 220.
(8) MACKINNON, L.: *ibid.*, 1955, **100**, p. 2.
(9) TRUELL, R., and BRONZO, J.: *ibid.*, 1953, **90**, p. 152.
(10) MASON, W. P.: *Bell System Technical Journal*, 1955, **34**, p. 903.
(11) MASON, W. P., and MCSKIMIN, H. J.: *Journal of the Acoustical Society of America*, 1947, **19**, p. 464.
(12) MASON, W. P.: 'Electromechanical Transducers and Wave Filters' (Van Nostrand, 1948), p. 81.
(13) Reference 12, p. 403.

(8) APPENDICES

(8.1) Reflection Coefficient

From eqn. (2) the terminating impedance is

$$Z'_F = jZ_F \frac{Z_T \tan\theta_T + Z_F \tan\theta_F}{Z_F - Z_T \tan\theta_T \tan\theta_F}$$

and is plotted in Fig. 7.

Let $$\frac{Z'_F}{Z_S} = ja \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where $$a = \frac{Z_F}{Z_S}\frac{Z_T \tan\theta_T + Z_F \tan\theta_F}{Z_F - Z_T \tan\theta_T \tan\theta_F}$$

The reflection coefficient is given by

$$R = \frac{Z'_F - Z_S}{Z'_F + Z_S}$$

Hence, substituting for Z'_F from eqn. (6),

$$R = \frac{ja - 1}{ja + 1}$$

$$= -(1 - ja)(1 + ja)^{-1}$$

And $$|R| = 1$$

while $$\phi_R = \pi - 2 \arctan a$$

(8.2) Integrating Effect of Crystal

It is required to find what output voltage occurs when a wavefront of known pressure and phase distribution strikes the crystal.

Assuming that the crystal is large enough for small areas to act independently of one another, the output voltage V is

$$V \propto \int_{r=0}^{r=a_1} p(r)\varepsilon^{j\phi(r)} 2\pi r dr$$

where r = Radius at which the elemental area $2\pi r dr$ is considered.

$p(r)$ = Pressure at r.
$\phi(r)$ = Phase of the pressure at that point (relative to some arbitrary zero).
a_1 = Radius of the crystal.

In order to simplify matters two assumptions are made.

(a) *Phase shift is proportional to radius.*

$$\phi(r) = Kr$$

K is a constant such that the net phase difference between centre and perimeter $= Ka_1$

(b) *Pressure p varies as 1/r.*

V is then proportional to $2\pi \int_0^{a_1} \frac{1}{r} r \varepsilon^{jKr} dr$

which gives $|E| \propto \frac{1}{K} \sin\frac{Ka_1}{2}$

Zeros of this function occur when $Ka_1 = 2n\pi$ (n = 1, 2, 3, etc.)

It is seen that the first zero occurs when there is 360° phase difference between that part of the wavefront arriving at the centre of the crystal and that part arriving at the outside. The input voltage is then zero.

(8.3) Effect of a Coupling Film on the Received Signal: Phase Effect

The phase shift occurring when a signal is transmitted through the coupling film from specimen to transducer may be most easily found by considering Fig. 6.

Applying transmission-line theory,

$$V \propto \frac{1}{j\omega C_0}(v_{T1} - v_{T2})$$

and $$v_{T1} = \frac{p_{0c} Z'_F}{Z'_F + Z_S}\ \frac{1}{jZ_F \sin\theta_F + jZ_T \tan\theta_T \cos\theta_F}$$

where p_{0c} = 'open-circuit' pressure, i.e. the pressure that would exist between the points FF_1 if no connection were made to these terminals.

This gives

$$V \propto \frac{1}{j\omega C_0}\left(1 - \frac{1}{\cos\theta_T}\right)\frac{1}{j(Z_F \sin\theta_F + Z_T \tan\theta_T \cos\theta_F)}\ \frac{1}{1 + \frac{Z_S}{Z'_F}}$$

Interest centres on the phase variation, given by the last term:

$$Z'_F = jaZ_S \text{ [from eqn. (6)]}$$

Hence the phase variation is given by the term

$$\frac{1}{1 + \frac{1}{ja}}$$

and the phase shift, ϕ_F, caused by the film, on reception is

$$\phi_F = \arc \tan (-a) = \frac{\phi_R - \pi}{2} \text{ (from Section 8.1).}$$

It may be shown that the same phase shift occurs when transmitting.

(8.4) Effect of a Coupling Film on the Transmitted Signal: Amplitude Effect

It is necessary to examine the change in the pressure amplitude in the rod as the thickness of the coupling film is altered. For simplicity the approximate equivalent circuit of Fig. 5 is used. The complete equivalent circuit may be drawn as in Fig. 12

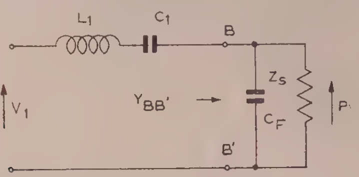


Fig. 12.—Approximate equivalent circuit for calculation of pressure amplitude at the face of the specimen.
For method of derivation of C_F, see Appendix 8.4.

(the capacitance $C_E/4$ may be neglected as it is more than 100 times greater than C_1).

Operation of a 1 cm² 10 Mc/s transducer near its 7th harmonic results in the following values for the circuit elements:

$$C_1 = 1{\cdot}4 \times 10^{-16} \text{ farad}$$
$$L_1 = 0{\cdot}04 \text{ henry}$$
$$Z_S = 15{\cdot}1 \times 10^5 \text{ ohms}$$

C_F is determined by the thickness of coupling film: values for it have been obtained by considering the true input admittance (Y_{BB}') to the right of BB′, as given by the line equation for a resistive load Z_S seen through a length of transmission line, l_F, Z_F, representing the film.

$$Y_{BB'} = \frac{1}{Z_F}\frac{Z_F + jZ_S \tan \theta_F}{Z_S + jZ_F \tan \theta_F}$$

or, in practice, a Smith chart may be used to simplify calculations.

This admittance may then be represented approximately by a shunt capacitance C_F across Z_S. Calculations give, for 0°, 5°, 10° and 20° films respectively, $C_F = 0$, 14×10^{-16}, 28×10^{-16}, and $56 \times 10^{-16} F$ or ratios of C_F/C_1 of 0, 10, 20, 40.

An equation may be derived relating the output pressure p across Z_S to the input voltage V_1:

$$p \propto \frac{1}{\left(j\omega L_1 + \frac{1}{j\omega C_1} + \frac{1}{\frac{1}{Z_S} + j\omega C_F}\right)} \frac{1}{\frac{1}{Z_S} + j\omega C_F}$$

$$\text{or} \quad |p| \propto \frac{1}{\sqrt{\left\{\left[1 + \frac{C_F}{C_1}\left(1 - \frac{f^2}{f_0^2}\right)\right]^2 + \left[\frac{1}{\omega C_1 Z_S}\left(1 - \frac{f^2}{f_0^2}\right)\right]^2\right\}}}$$

where f_0 is the harmonic of the resonant frequency. This is plotted in Fig. 10.

An almost identical expression results if the relation between output voltage and applied pressure is determined.

621.373.2.029.6

Paper No. 2185 R
Nov. 1956

RESEARCHES INTO SPARK GENERATION OF MICROWAVES

By M. H. N. POTOK, B.Sc., Ph.D., Associate Member.

(*The paper was first received* 14*th May, and in revised form* 11*th July*, 1956.)

SUMMARY

A study of microwave spark generators shows that they have certain useful characteristics. In particular, the wide band of frequencies generated permits the use of filters to select any desired band of any width. Both post and iris type of filters were used. The measurements of wavelength and bandwidth were carried out with the help of a Boltzmann interferometer using a cathode-ray oscilloscope and a camera for recording purposes. Very satisfactory results were obtained in the 2 to 4 cm region and less so down to 8 mm.

LIST OF SYMBOLS

$g(\omega)$ = Frequency spectrum.
$f_0 = \omega_0/2\pi$ = Natural frequency of radiation of a dipole.
α = Attenuation = δf_0.
δ = Logarithmic decrement of radiation.
$f_c = \omega_c/2\pi$ = Waveguide cut-off frequency.
s = Mirror separation in the interferometer.
c = Velocity of electromagnetic waves.
$\tau = 2s/c$ = Time displacement between waves reflected by the two mirrors of the interferometer.

(1) INTRODUCTION

Recent developments in microwaves have revived the interest in spark generators as possible sources of microwave energy in the millimetre region. A review of past researches into spark generators has been given by the author elsewhere.[1] Further research in that direction in this country[2] as well as in the United States,[1] Russia,[1] Germany[1] and Japan[3] have led to somewhat pessimistic conclusions, as far as power output is concerned, especially in view of the development of the backward travelling-wave tube.[4]

Nevertheless, the microwave spark generator has certain characteristics and advantages which deserve further examination. These stem from the fact that the coherent generated wave has a wide spectrum, being unique in this respect among the common sources of microwaves. The spectral characteristic is caused by the spark-generated wave being a damped sinusoid of the general form $\varepsilon^{-\alpha t}\sin\omega_0 t$ whose power spectrum is given by

$$|g(\omega)|^2 = \frac{\omega_0^2}{4\pi^2(\omega-\omega_0-j\alpha)(\omega-\omega_0+j\alpha)(\omega+\omega_0-j\alpha)(\omega+\omega_0+j\alpha)}$$

This is shown in Fig. 1 for various values of the logarithmic decrement $\delta = \alpha/f_0$. This damped oscillation is radiated when a high-voltage pulse is applied across a dipole consisting of two very short cylinders or spheres suspended in a liquid of high breakdown strength, such as oil. The radiated wave has a natural wavelength, λ_0, of between two and five times the overall dimension of the dipole (depending on the ratio of cross-section to length and on the method of suspension of the dipole in liquid), while the logarithmic decrement of the wave is of the order of 0·5 to 1·0. Fig. 2 shows the relation between the ratio of wavelength to dipole length and the ratio of dipole length to diameter obtained by many workers over the past 50 years.[1, 5]

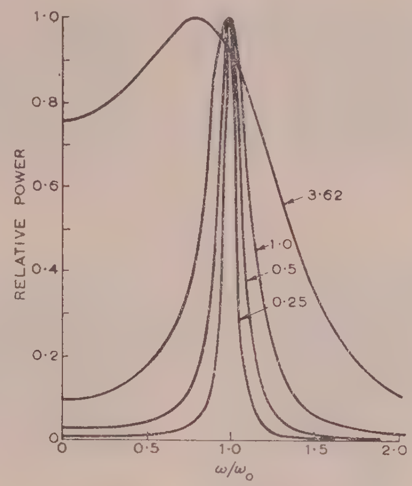


Fig. 1.—Spectrum of radiating, self-oscillating dipoles of various logarithmic decrements.

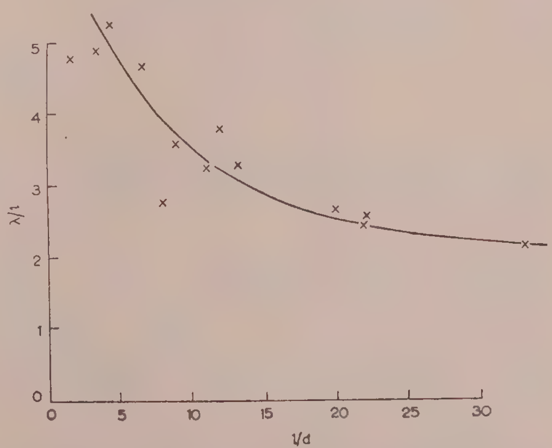


Fig. 2.—The relation between the ratio of the natural wavelength to the physical length (λ_0/l), and the ratio of the length to the diameter of dipoles (l/d), as found by several workers over the last 50 years.

The equipment required to produce such waves is comparatively simple, which presents another important aspect of microwave spark generation.

(2) EQUIPMENT

The basic equipment consists of a source of high voltage and the dipole. The author has experimented with several types of

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Dr. Potok is in the Department of Electrical Engineering, Royal College of Science and Technology, Glasgow.

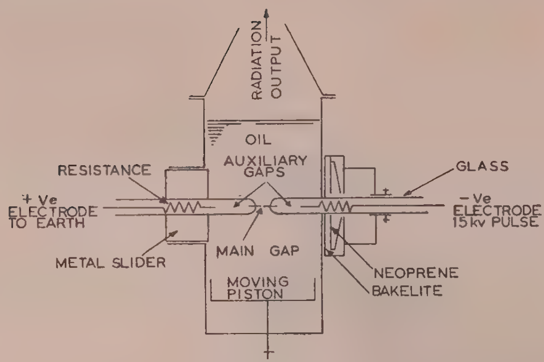


Fig. 3.—Schematic of the basic radiator in a microwave spark generator.

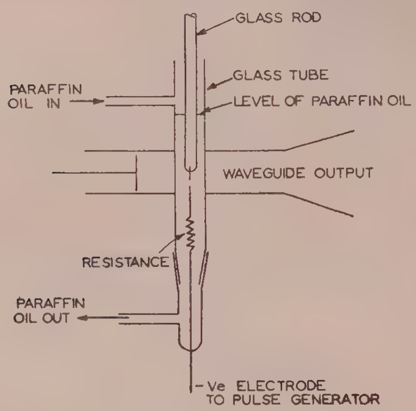


Fig. 4.—A simplified version of the radiator.

high-voltage supplies. The experience gained indicates that, while even sinusoidally varying voltage of the order of 10 kV will produce microwave radiation from the dipoles, the power fluctuates considerably unless a pulse generator is used producing pulses with a rise time better than 1 microsec. The dipoles must, of course, be suspended in oil or similar liquid. The oil-gap between the two elements (Fig. 3) constituting the dipole must be very small indeed. On the other hand, the two auxiliary gaps between the dipole and the high-voltage electrodes should be in air, one being made as large as possible. This results in a much quieter discharge and much more regular microwave output—an effect which does not appear to have been commented on in the published work, and yet is quite pronounced, leading to a great improvement in working and in stability, which is most important if measurements are to be carried out. Further work has shown that very good results can be obtained by dispensing with one half of the dipole and the earth return. This allows a simplified construction of the radiator system, which can now be mounted direct across a waveguide as in Fig. 4. The oil container is made in two parts for ease of assembly, and oil can be circulated if required (constant sparking between the radiating element and the h.v. electrode causes the oil to become polluted within an hour or so, which, while not harmful at once, leads to an eventual drop in power).

For best results the h.v. electrode has to be made sharp and the radiating element has to be brought into very close proximity to its tip. For this purpose the glass or polystyrene rod holding the radiating element should be held so as to allow fine adjustment in three dimensions.

The pulse generator used for the experiment is a conventional hard-valve type, using a 'boot-strap' circuit to drive the type 715 B tetrode which, when cut off, holds off a direct voltage of 20 kV obtained, by rectifying in a Cockcroft–Walton quadrupler, a 2000 c/s supply. Fig. 5 shows the complete circuit.

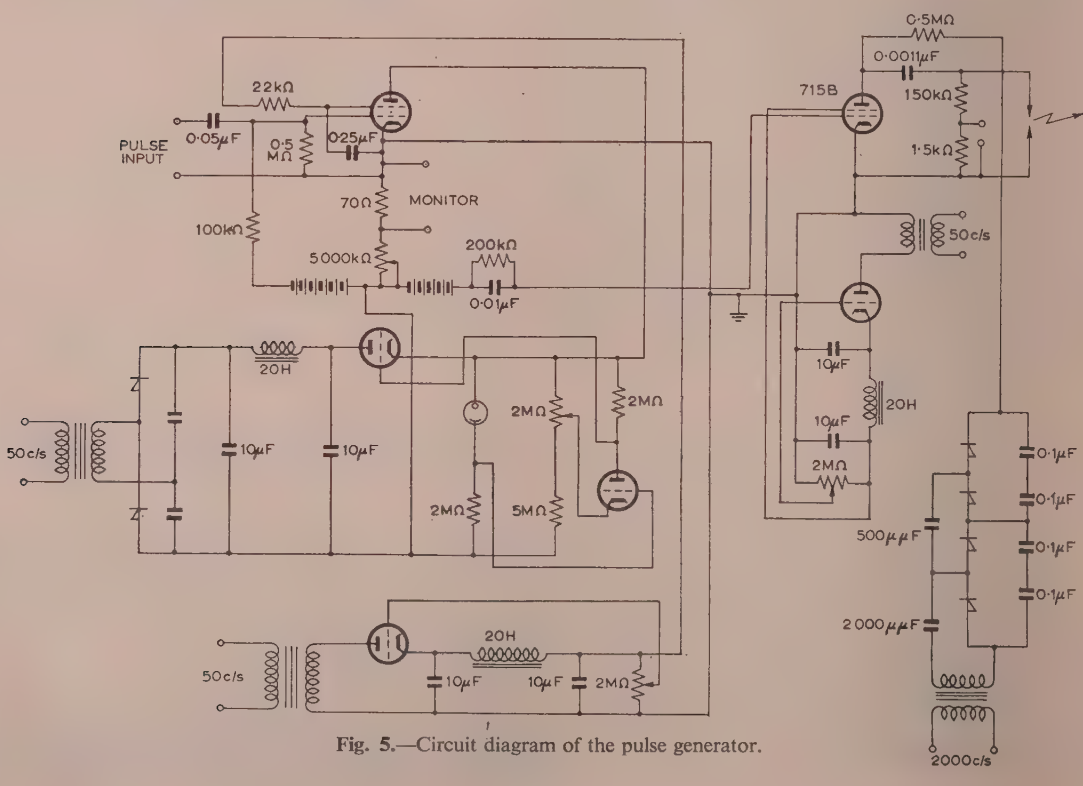


Fig. 5.—Circuit diagram of the pulse generator.

(3) MEASUREMENT TECHNIQUE

All measurements have been obtained by the medium of a Boltzmann interferometer employing a photographic recording technique which is particularly suitable for this purpose.

(3.1) Recording

Since the power radiated by the spark generator tends to fluctuate somewhat even with most careful adjustment of radiator, a method commonly adopted to measure the output is to use an instrument with a long time-constant. This method was discarded, because of its low sensitivity and slow operation, in favour of a method novel in this connection using a cathode-ray tube and a camera.

Since the power is radiated in the form of pulses at a fixed frequency of the driver, the receiver (usually a crystal) feeds into a narrow-band amplifier tuned to the driver frequency. The output is applied to the Y plates of the oscilloscope. The resultant vertical line is proportional to the power received by the crystal. This is then photographed by exposing the film for a fixed time, the resultant trace integrating the detected power.

If now either the film or the beam is traversed horizontally, moving at a constant speed, a band will appear on the film whose edges are somewhat blurred owing to the power fluctuation. By tracing along the edge a line keeping to a constant brightness (which requires a little skill, but can be done automatically by using a suitable photometer), the plot of power is obtained against a variable parameter which controls the movement of film or beam. All the results quoted here have been obtained by this method. Some typical interferograms are shown in Fig. 6.

(3.2) The Boltzmann Interferometer

The theory of the Boltzmann interferometer has been developed by Brown and Farrands elsewhere.[6] It has been shown by them that, if a wave whose spectrum is $g(\omega)$ is reflected by a Boltzmann interferometer, the total relative reflected power is given by the expression

$$P = \int_0^\infty |g(\omega)|^2 d\omega + \int_0^\infty |g(\omega)|^2 \cos \omega\tau d\omega$$

where $\omega\tau$ is the total phase shift between the waves reflected by the two mirrors. If the angles of incidence and reflection are small, $\tau = 2s/c$, where the mirror separation is s and the velocity of propagation of the wave is c.

The second term in the expression for P is seen to be a cosine transform of $|g(\omega)|^2$, and hence by taking an inverse cosine transform,

$$|g(\omega)|^2 = \int_0^\infty \left[\int_0^\infty |g(\omega)|^2 \cos \omega\tau dw\right] \cos \omega\tau d\tau$$

$$= \int_0^\infty \left[P - \int_0^\infty |g(\omega)|^2 d\omega\right] \cos \omega\tau d\tau$$

The expression in the square bracket is clearly the difference between the total power at any mirror separation and that at infinite mirror separation; hence $|g(\omega)|^2$ can be found by a suitable computation, when the interferogram is known.

This assumes that the frequency response of the receiver is uniform throughout the spectrum, which, of course, is not the case. On the one side waveguides act as high-pass filters, and on the other, the crystal sensitivity falls off rapidly as frequency is increased.[7] With comparatively narrow-band filters this is not

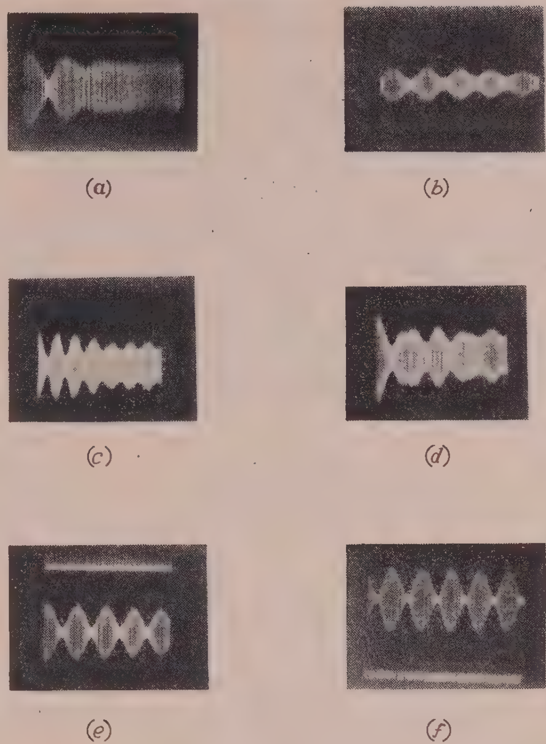


Fig. 6.—Examples of interferograms:

(a) Interferogram obtained without any filter.
(b) The wave filtered by a three-post filter, tuned to the natural frequency of the radiator.
(c) The wave filtered by a post filter, tuned to twice the natural frequency of the radiator.
(d) An interferogram obtained when the wave spectrum contains two pronounced peaks (as in Fig. 11, with two posts only).
(e) The interferogram obtained with an iris filter.
(f) The interferogram obtained with a 3 cm klystron.

significant as long as we keep well away from the cut-off frequency of the waveguide; but when the full generated spectrum is examined, or very wide band filters are used, the computed spectrum has to be divided by the known frequency response of the receiver.

The photographic technique described earlier is very useful for taking the interferograms. The sending and receiving horns are placed side by side, and the mirrors of the interferometer are adjusted to reflect equal powers. The receiver provides vertical deflection of the cathode-ray beam while the moving mirror, by gearing to a potentiometer, traverses the beam across the screen with the camera shutter open. Thus the film records the power-output variation against mirror displacement (interferogram). A probe in the sending horn feeding the second beam provides a monitor of power input should this vary during the run.

Typical interferograms obtained with a monochromatic source (klystron) and with a decaying oscillation are shown in Fig. 6.

The accuracy of the interferogram depends on balancing the mirrors, but it can be shown that, even if the two mirrors reflect slightly different proportions of power, say the ratio of reflected powers by the two mirrors is $1/(1 + a)$, the total reflected power is given by

$$P = \int_0^\infty |g(\omega)|^2[2(1 + a) + a^2]d\omega + \int_0^\infty |g(\omega)|^2 2(1 + a) \cos \omega\tau d\omega$$

If a is small, so that $2(1 + a) \gg a^2$, the only difference between

this and the perfectly balanced case is a constant factor $2(1 + a)$, which, since in any case the measured power output is relative, introduces no error.

When the cosine transform of the interferogram is computed, the first step is to draw the axis corresponding to the condition of infinite separation between mirrors, i.e. $\int_0^\infty |g(\omega)|^2 d\omega$. If the second beam is not used for monitoring, this axis can be drawn automatically by it, but more frequently, in particular, with spark generators giving slow changes in power level as well as pulse-to-pulse fluctuations, monitoring is essential and the axis has to be drawn in on the completed interferogram. In that case an error may be introduced owing to the wrong axis being chosen. This point was examined by computing the spectrum from an interferogram in which four different axes were assumed.

The results shown in Fig. 7 indicate that the error is not a serious one.

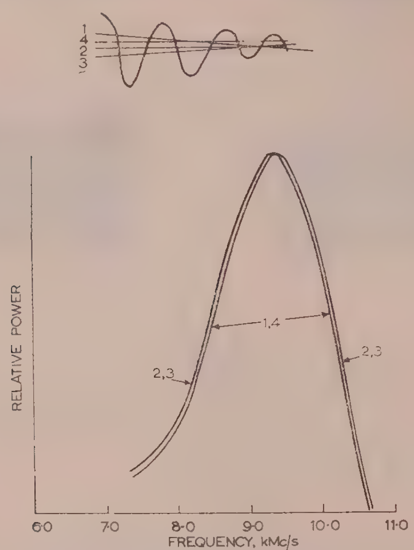


Fig. 7.—The effect of shifting the zero axis of the interferogram on the computed spectrum.

The difficulty of drawing in the axis is increased when the interferogram cannot be taken far enough to reach the constant power level corresponding to very large mirror separation, such as may be required when the decaying wave has a very low decrement. A computation of $|g(\omega)|^2$ based on an incomplete interferogram may introduce a serious error. Its magnitude may be gauged by reference to Fig. 8, which shows how the spectrum appears to spread out as the interferogram of an exponentially decaying wave is cut shorter and shorter. The necessary correction in calculating the half-power bandwidth of the spectrum can be estimated with the help of Fig. 9.

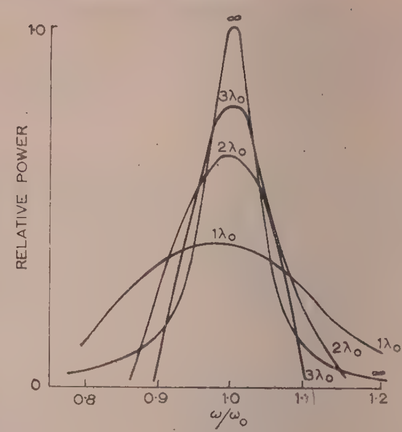


Fig. 8.—Spectrum of power computed from the Boltzmann interferogram, taking into account a length of $1\lambda_0$, $2\lambda_0$, $3\lambda_0$, ∞, when the logarithmic decrement of the analysed wave is $0{\cdot}25$.

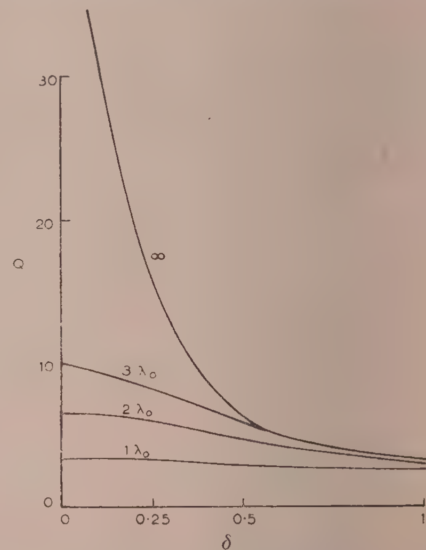


Fig. 9.—Correction curves for the Q-factor when only a part of the interferogram is available.

(3.3) The Standing-Wave Method

Instead of using a Boltzmann interferometer one fixed reflector can be used and the standing wave in front of it measured by a suitable probe. The power measured by an aperiodic probe would be given by

$$P = \int_0^\infty |g(\omega)|^2 d\omega - \int_0^\infty |g(\omega)|^2 \cos \omega\tau d\omega$$

where $\tau = 2x/c$ and x is the distance between the probe and the reflector. Again, an inverse transform of the second term on the right-hand side would give $|g(\omega)|^2$. This method has been used by Kawano,[3,8] but in the author's experience it involves much greater difficulties in setting up and avoiding diffraction errors than the Boltzmann-interferometer method.

(3.4) Waveguide Methods

The Boltzmann interferometer could be substituted by a hybrid-T with two short-circuiting plungers, one fixed and one moving. On the other hand, the standing-wave method could be applied by using an ordinary waveguide standing-wave-ratio meter with a terminating short-circuit. In either case the quantity previously designated $\omega\tau$ now becomes $\omega 2x/u_g$, where x is the difference between distances of the two plungers and the hybrid centre-line or the distance between the probe and short-circuit respectively, and u_g is the guide velocity of propagation.

Since $$u_g = \frac{\omega c}{\sqrt{(\omega^2 - \omega_c^2)}}$$

vhere ω_c is the cut-off angular frequency and c is the velocity of ropagation in unbounded medium

$$\omega\tau = \frac{2x}{c}\sqrt{(\omega^2 - \omega_c^2)} = \omega'\tau'$$

Hence the total detected power is

$$P = \int_0^\infty |g(\omega)|^2 d\omega \pm \int_0^\infty |g(\omega)|^2 \cos \omega'\tau' d\omega$$

nd, as before, the power spectrum $|g(\omega)|^2$ can be found from he cosine transform of the variable term after correcting for he difference between ω' and ω. The main disadvantage of his method is the need of a very long movement of either the robe or the short-circuiting plunger—much longer than that alled for in the Boltzmann interferometer, because, of course, $_g$ increases as ω approaches ω_c. The consequent technical ifficulty as well as the increased attenuation again favour the se of the Boltzmann interferometer.

(4) REDUCTION OF RECEIVED BANDWIDTH

Very few applications are known for which the whole radiated and is required. The reduction of the band can be obtained y placing the radiating dipoles in a cavity, by filtering the adiated wave or by dispersing the radiation by a suitable lens, rating or echellette. All three have been used with more or ess success, but experience indicates that the filter is the most ersatile. Among its advantages is the facility it provides for btaining several outputs split into bands of desired width and entre.

Simple waveguide filters such as the post and the iris type give xcellent results.

The use of filters with the Boltzmann interferometer illustrates lso the important application of spark microwave generators o the study of microwave devices. By suitable computations erformed on the interferograms obtained first without and hen with the component in circuit, its frequency characteristic an be obtained over any desired frequency band over the whole nicrowave range.

(4.1) Experiments with Post Filters

As is well known, two posts parallel to the E vector almost alf a wavelength apart along the centre-line of the guide in he direction of propagation act as a band-pass filter whose atio of centre frequency to half-power bandwidth—referred to ereafter as the Q-factor of the filter—is of the order of 10 to 100. f a stub of variable length is placed half-way between the posts he filter can be tuned by it.

Using a radiator whose natural wavelength lies around 3·4 cm in air) attempts have been made to find the relation between ilter centre frequency and the separation of the posts made of $\frac{1}{16}$ in-diameter brass in a standard X-band waveguide (0·9 in × ·4 in). The results are given in Fig. 10. The computations eading to these results have been carried out with the help of he Beevers–Lipson 3° strips.[9, 10] This method of Fourier nalysis is widely applied by the crystallographers and appears o be less well known to engineers than it deserves; it provides, hort of a computer, the quickest method of computing transorms (see also Appendix). Although Fig. 10 does not show it, s the posts are brought closer together, the output contains an ver-increasing component of band centred on the natural requency of the radiator, so that the filter ceases to be useful vith this radiator when the posts are less than 8 mm apart. A onsiderable improvement is obtained by using three or more quidistant posts as seen in Fig. 11. With three posts 8 mm apart and two 6 B.A. tuning screws half-way between them, the relation between depth of tuning screws and frequency is that given in Fig. 12, which also shows the relative power in the transmitted band. Thus this arrangement permits a band centred anywhere between 8 and 13·75Gc/s to be selected, the

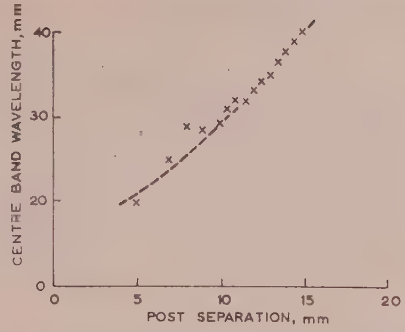


Fig. 10.—Wavelength at the centre of the filtered band versus separation of $\frac{1}{16}$ in posts in an X-band waveguide.

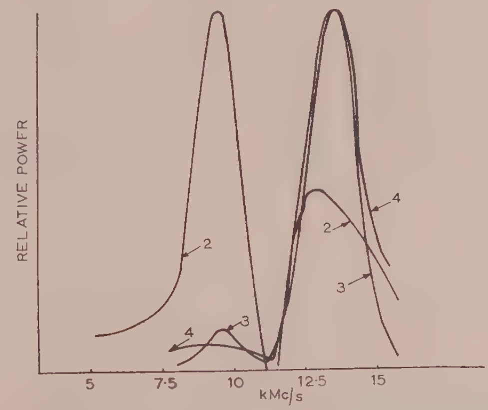


Fig. 11.—Power spectrum of an exponentially decaying wave passed through a post filter consisting of 2, 3 and 4, 6 B.A. posts 8 mm apart from one another along the centre-line of an X-band waveguide.

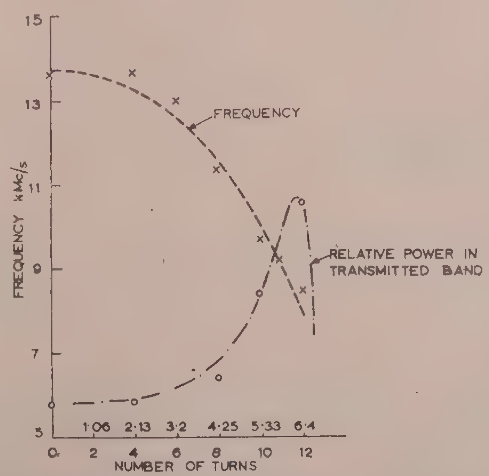


Fig. 12.—Frequency at the centre of the band, and the relative power in the band at various depths of the tuning screw.

Q-factor of the filter being everywhere of the order of 20, i.e. the bandwidth is of the order of 0·5 Gc/s.

(4.2) Experiments with Iris-Type Filter

By placing two inductive irises almost half a wavelength apart across a guide, high Q-factor filters can be obtained. The filter used had an iris width of 0·2 in, and its length was $\frac{13}{32}$ in, in a standard X-band waveguide. The calculated Q-factor of this filter is 90, centred on 2 cm wavelength. A central 6 B.A. screw was used for tuning. The results obtained when the depth of tuning screw was varied appear in Fig. 13. The bandwidth was

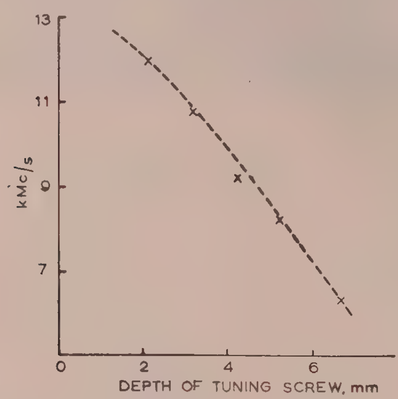


Fig. 13.—Centre-band frequency versus the depth of the tuning screw for the iris filter.

of the order of 100–250 Mc/s. It was not possible to estimate this with any degree of accuracy because the decrement was so low that the interferogram required for accurate computation would have to be longer than the equipment allowed.

(4.3) Experiments at Millimetre Wavelength

The experiments were progressively extended to shorter wavelengths, but because of the technical difficulties of making filters in Q-band guides, and also in view of the rapidly decreasing powers, results were less satisfactory. However, quite good results were obtained by using tapered guide sections to raise the cut-off frequency. By these means bands of an estimated width of 1–2 Gc/s centred anywhere between 2 cm and 8 mm could be obtained. Another method of varying the wavelength was to insert various dielectrics into the Q-band guide at the receiver end, giving entirely satisfactory results at a few frequencies. A complete range of dielectric constants—such as can be obtained by varying the density of polystyrene foam—would be of great value here.

(5) CONCLUSIONS

The experiments have established a means of producing wide-band microwave generators of any desired bandwidth and centred anywhere between 8 mm and 4 cm. Comparatively simple, cheap and reliable equipment is required giving a stable source which can be used for demonstration and testing.

The source when used in conjunction with a Boltzmann interferometer, following the cathode-ray photography technique described, can be used to study the characteristics of wide-band waveguide components over the whole spectrum.

It would appear very advantageous to make the Boltzmann interferometer work direct into some form of a simple computer which would produce the frequency spectrum direct, as the mirrors move apart. This work is now in progress, and when complete, it is expected to give a very versatile tool for the analysis of frequency characteristics of microwave devices.

The power generated is very small, of the order of 10–20 μW and falling off rapidly as the frequency is increased, but with a well-designed pulse generator producing sharp pulses, the power can be increased by raising the repetition frequency. Linear increase in power has been observed up to 5 000 c/s, which was all the equipment could cope with, but it indicates that higher repetition frequency could be used. Also, the pulsed nature of radiation should allow the development of coherent detectors able to make full use of the small powers available.

(6) ACKNOWLEDGMENT

The author wishes to thank Prof. F. M. Bruce for providing the facilities for carrying out the research at the Royal College of Science and Technology, Glasgow.

(7) REFERENCES

(1) Potok, M. H. N.: 'A Critical Review of Researches into Millimetric Wave Spark Generators', *Journal of the British Institution of Radio Engineers*, 1953, **13**, p. 490.
(2) Farrands, J. L.: 'The Generation of Millimetre Waves', *Proceedings I.E.E.*, Monograph No. 112 R, December, 1954 (**102 C**, p. 98).
(3) Tadasu Kawano: 'Research Reports on Spark Generated Microwaves', *Bulletins of the Faculty of Engineering, University of Hiroshima*, 1952–55.
(4) Pierce, J. R.: 'Some Recent Advances in Microwave Tubes', *Proceedings of the Institute of Radio Engineers*, 1954, **42**, p. 1735.
(5) Lindman, K. F.: Research Reports, *Acta Academiae Aboensis*, 1929–47.
(6) Farrands, J. L., and Brown, J.: 'Boltzmann Interferometer', *Wireless Engineer*, 1954, **31**, p. 81.
(7) Torrey, H. C., and Whitmer, C. H.: 'Crystal Rectifiers' (McGraw-Hill, New York, 1948).
(8) Tadasu Kawano: 'Propagation of Heavily Damped Microwaves through a Waveguide', *Journal of the Institute of Electrical Communication Engineers of Japan*, 1953, **36**, p. 591.
(9) Beevers, C. A.: 'Fourier Strips at 3° Interval', *Acta Crystallographica*, 1952, **5**, p. 670.
(10) Lipson, H., and Cochran, W.: 'The Determination of Crystal Structures' (G. Bell and Sons, 1953).
(11) Twiss, R. Q.: 'On the Generation of Millemetre Radiation', *S.E.R.L. Technical Journal*, 1952, **2**, p. 10.

(8) APPENDIX

(8.1) A Note on Beevers–Lipson Strips

The computation of the Fourier cosine transform of a function $f(\tau)$ involves finding the area under the curve $f(\tau)\cos\omega\tau$ for all relevant values of ω. The simplest method of doing so is to find the value of the function $f(\tau)$ at suitable intervals τ' and to compute the sum

$$\sum_{n=0}^{\infty} f(n\tau')\cos\omega n\tau'$$

for a number of values of ω. Since the function $f(\tau)$ converges fairly quickly to 0, this involves taking some 30 to 60 ordinates (the number of ordinates required increases with the Q-factor of the filter).

The computation can be accelerated by drawing up the following Table:

$f(0)\cos 0$ $f(\tau')\cos 0$ $f(2\tau')\cos 0$ $f(3\tau')\cos 0$ . . . $f(n\tau)\cos 0$	$f(0)\cos 0$ $f(\tau')\cos \omega'\tau'$ $f(2\tau')\cos 2\omega'\tau'$ $f(3\tau')\cos 3\omega'\tau'$ . . . $f(n\tau')\cos n\omega'\tau'$	$f(0)\cos 0$ $f(\tau')\cos 2\omega'\tau'$ $f(2\tau')\cos 4\omega'\tau'$ $f(3\tau')\cos 6\omega'\tau'$ . . . $f(n\tau')\cos 2n\omega'\tau'$	$f(0)\cos 0$ $f(\tau')\cos 3\omega'\tau'$ $f(2\tau')\cos 6\omega'\tau'$ $f(3\tau')\cos 9\omega'\tau'$ . . . $f(n\tau')\cos 3n\omega'\tau'$	....
$\sum_{n=0}^{n} =$	relative amplitudes of the cosine transform at: ω'	$2\omega'$	$3\omega'$	....

The rows of the above Table are, of course, the Beevers–Lipson strips, available for all values of $f(\tau)$ from $+100$ to -100 (with extension to $\pm 1\,000$) for 30 ordinates, the interval between the columns being 3°. With the help of these, the complete computation takes less than 20 min and can be further speeded up, since not all columns have to be summed. Thus, if it is enough to know the transform every 0·5 Gc/s between 8 and 12 Gc/s then, by setting $\omega' = 10^9\pi$, only the 16th to the 24th columns need be considered, in which case, the interferogram having been read every 2·5 mm of the mirrors' displacement, the assembly of the Table and the addition can be completed in 10 min.

The interval between the columns (θ^0) defines the product $\omega'\tau'$, and since $\tau' = 2s'/c$ and $\omega' = 2\pi f'$, we have $f's' = \theta c/720$, where s' represents the interval at which the interferogram has to be read, and this should be chosen to be as large as permissible to save computation time.

If the available 3° Beevers–Lipson strips give too coarse a result, a new set of strips has to be made either by direct calculation or by interpolating the existing ones. Also, more than 30 ordinates may be required. The making of these strips involves less work than might at first appear, because the transform will be sought in a comparatively narrow band; hence only a part of each strip has to be filled.

DISCUSSION ON 'THE LAUNCHING OF A PLANE SURFACE WAVE'*

Prof. R. H. DuHamel (*Illinois, U.S.A.: communicated*): This concerns the discrepancy between the theoretical and measured values of launching efficiency for a slot source above a dielectric coated plane conductor.

Let us represent the transition from the input waveguide to

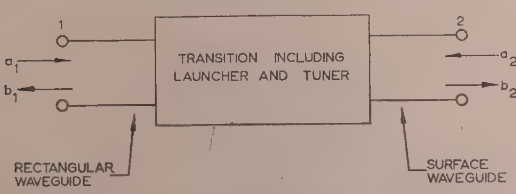


Fig. B.—Equivalent circuit of launcher and connecting waveguides.

the dielectric-coated surface by the two-port network shown in Fig. B. The scattering matrix† representation of the transition is

$$\left.\begin{aligned} b_1 &= S_{11}a_1 + S_{12}a_2 \\ b_2 &= S_{21}a_1 + S_{22}a_2 \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (A)$$

The launching efficiency is defined by

$$\eta_{12} = \frac{|b_2|^2}{|a_1|^2 - |b_1|^2} = \frac{|S_{12}|^2}{1 - |S_{11}|^2} \quad . \quad . \quad . \quad (B)$$

and the collecting efficiency by

$$\eta_{21} = \frac{|b_1|^2}{|a_2|^2 - |b_2|^2} = \frac{|S_{12}|^2}{1 - |S_{22}|^2} \quad . \quad . \quad . \quad (C)$$

Mr. Rich stated that the launching efficiency was equal to the reflection coefficient as measured in the rectangular waveguide with a short-circuit plate on the surface waveguide after first adjusting the transition so that $S_{11} = 0$. Denoting the reflection coefficient of the short-circuit by $\varepsilon^{j\beta}$, a simple calculation of the reflection coefficient yields

$$\frac{b_1}{a_1} = \frac{S_{12}^2\varepsilon^{j\beta}}{1 - S_{22}\varepsilon^{j\beta}} \quad . \quad . \quad . \quad . \quad . \quad (D)$$

which is equal to the launching efficiency only if the launcher is matched to both the input and output waveguides, i.e. $S_{11} = S_{22} = 0$. Since it appears that no precautions were taken to set $S_{22} = 0$, and since without proper tuning this will be the exception rather than the rule, it is likely that the efficiency measurements were in error, especially for the slot sources, since

* Rich, G. J.: Paper No. 1783 R, March, 1955 (see 102 B, p. 237).

† For further details see Montgomery, C. G., Dicke, R. H., and Purcell, E. M.: 'Principles of Microwave Circuits', Radiation Laboratory Series No. 8 (McGraw-Hill, 1948).

they are poorly matched to the surface waveguide. The reciprocity theorem may be used to justify his method of measurement only if the system is completely matched.

The launching efficiency may be determined correctly by measuring the scattering coefficients by Deschamps's method.*

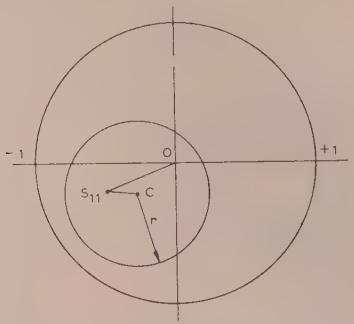


Fig. C.—Determination of scattering coefficients from image circle and point S_{11}.

$$|S_{11}| = \overline{OS_{11}}$$
$$|S_{22}| = \overline{S_{11}C}/r$$
$$|S_{12}|^2 = r[1 - |S_{22}|^2]$$

Referring to Fig. C, it is only necessary to determine the image circle and the point S_{11}. Eqns. (B) and (C) may be evaluated by using the formulae in Fig. C.

A comparison of Mr. Rich's results with Cullen's† predicted launching efficiency of 95% for a slot source means little, since Cullen's calculations were for $\lambda_g/\lambda_0 = 0 \cdot 9$, whereas Mr. Rich's measured value of λ_g/λ_0 was $0 \cdot 978$. The theory would most likely predict a maximum launching efficiency considerably less than 95% for the latter case.

* Storer, J. E., Shenigold, L. S., and Stein, S.: 'A Simple Graphical Analysis of a Two-Port Waveguide Junction', *Proceedings of the Institute of Radio Engineers*, 1953, **41**, p. 1004.

† Cullen, A. L.: 'The Excitation of Plane Surface Waves', *Proceedings I.E.E.*, Monograph No. 93 R, February, 1954 (101, Part IV, p. 225).

Mr. G. J. Rich (*in reply*): While agreeing in general with Prof. DuHamel's treatment of the launching problem, I would be inclined to use a different definition of η from the one he has used. I would put

$$\eta_{12} = \frac{|b_2|^2}{|a_1|^2} \text{ and } \eta_{21} = \frac{|b_1|^2}{|a_2|^2} \quad . \quad . \quad . \quad . \quad (E)$$

i.e. the launched or collected power is compared with the incident power, not with the net incident power ($|a_1|^2 - |b_1|^2$ or $|a_2|^2 - |b_2|^2$). The above definitions lead to

$$\eta_{12} = \eta_{21} = S_{12}^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (F)$$

if both the surface and the generator are matched. The launching efficiency so defined is equal to the collecting efficiency.

I do accept Prof. DuHamel's conclusion that the reflection coefficient of a short-circuited surface as seen from the input waveguide is not the same as η unless $S_{22} = 0$, and I am grateful to him for pointing this out and for drawing my attention to Deschamps's method. I agree that a slot source is but poorly matched to an incoming surface wave. Analysis by Deschamps's method of the results of a recent experiment with a flare-type launcher yields the following approximate results:

$$\left.\begin{aligned} |S_{11}| &= 0 \cdot 052 \\ |S_{12}| &= 0 \cdot 70 \\ |S_{22}| &= 0 \cdot 05 \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (G)$$

S_{22} is thus quite small for this type of launcher, and within the limits of experimental accuracy the values of η obtained from the scattering matrix by either of the definitions are the same as those given by the less rigorous method of taking η as being equal to the mean reflection coefficient.

INDEX TO VOLUME 103, PART B

1956

ABBREVIATIONS

(P)—Address, lecture or paper.
(p)—Subject dealt with in a paper or address.
(D)—Discussion on a paper.
(A)—Abstract of paper or address.

G

H

I

J

K

L

M

R

S

T

UNWIN BROTHERS LIMITED, WOKING AND LONDON

PROCEEDINGS OF THE INSTITUTION OF ELECTRICAL ENGINEERS

Part B. RADIO AND ELECTRONIC ENGINEERING (INCLUDING COMMUNICATION ENGINEERING) NOVEMBER 1956

CONTENTS

Bibliographical References.—It is requested that bibliographical reference to an Institution paper should always include the serial number of the paper and the month and year of publication, which will be found at the top right-hand corner of the first page of the paper. This information should precede the reference to the Volume and Part.
Example.—SMITH, J.: "Reflections from the Ionosphere," *Proceedings I.E.E.*, Paper No. 3001 R, December, 1954 (102 B, p. 1234).

Published by The Institution, Savoy Place, London, W.C.2. *Telephone:* Temple Bar 7676. *Telegrams:* 'Voltampere, Phone, London.'
Printed by Unwin Brothers Limited, Woking and London.